How Technologies Work

An Introductory Course

by

Rajesh R. Parwani

First Edition : July 2002
Second Edition: Jan 2003.
Third Edition : July 2003.
Fourth Edition : Jan 2005.

Preface

These are unedited outlines of the lectures conducted between 2001 and 2010 at USP/NUS, Singapore. Additional discussion and reading assignments supplemented these. The exercises at the end of each chapter are the basis for the tutorial questions and are meant in many cases to encourage the students to take an active part in the learning process by searching for answers through reading, thinking, discussion and experimentation.

Acknowledgements
The figures were drawn by David Phang Lip Pin while Karunanethy Kalaivani assisted with the summary of supplementary readings. They were both students who once took this module.

Contents

Chapter 1

Introduction

If it's green, it's biology, If it stinks, it's chemistry, If it has numbers it's math, If it doesn't work, it's technology
–Author Unknown

1.1 Overview

- **What is this module about?**
  This module is about technology: How human-made products work, how they are constrained by science and other factors, and how technology impacts on society and the environment. The lectures will cover mostly the first part, while the exercises and project work will enable you to explore the other two aspects.

- What is **technology**?
  The word originates from the Greek word *tecne* meaning skill, and the word *logia* meaning study or science. We will use 'technology' to refer both to human-made products and also sometimes to the art of creating those products.

  Science constrains what technology is able to achieve, for regardless of our ingenuity we cannot devise products that defy the laws of physics (Nature). However as we shall see, science is not the only constrain on technology, and furthermore, many wonderful technologies, like the guitar, were invented, or evolved with little explicit use of science.

- What is Science and the Scientific Method ?
  One may say that Science is the study of Nature, or the study of what exists. A scientist seeks to find the most economical and accurate description of natural phenomena, using the scientific method of testing hypothesis and models against experiment. Once a model has been well tested and is of sufficient generality, it is called a theory. Theories in the physical sciences tend to be mathematical, probably

because mathematics provides a relatively unambiguous and "culture free" language for the description of the relationships between quantitative data.

- What is engineering?
  The word "Engineer" originates in the Latin *ingeniatorem*, referring to a person who is ingenious at devising. Engineering is the art of creating useful products that do not exist naturally, or in enhancing those which do. An engineer follows the scientific method but the designs he makes are constrained by other factors as well. Indeed engineering often requires the search for an optimal solution, or compromise, between performance, cost, materials (weight, durability etc), politics, aesthetics, ethics and profitability. Furthermore, similar problems in different cultures can give rise to different solutions. Thus engineering often sythesises scientific knowledge with that from the humanities.

- What is the objective of this course ?
  The aim of the course, **How Technologies Work**, is to develop a keener sense of awareness in students of the technologies that shape our modern world. The students will study
  (1) The major and basic physical principles behind the operation of various technologies, and how those principles constrain the operation, size or capabilities of products.
  (2) In addition, in the reading assignments and the group project, the students will study and debate some important scientific and sociological issues relating to the development and evolution of technologies, and their impact on society.
  (3) This is not a module on engineering. Neither is this a physics module. Rather, it is somewhat closer to a study in reverse-engineering.

1.2 Technology Lock-In: The Typewriter Keyboard

When typewriters were first invented, there was no industry standard and each manufacturer was free to come up with his own keyboard layout. One of the first patents for a typewriter was filed in 1867 by Chirstorpher Latham Sholes who designed the QWERTY keyboard we are all familiar with, so called after the first six letters of the keyboard arrangement. The reason behind this particular arrangement is that this was the most inefficient one that Sholes could come up with! Early typewriters suffered from a design problem where the keybars that typed the letter onto the paper often jammed when typists typed too quickly. So instead of correcting this problem, Sholes decided on slowing down typists by placing the most frequently used words on the periphery of the keyboard so that the typist's fingers have

to travel further.

Even though many other more efficient designs were introduced in the 1880s, the QWERTY keyboard, having the first mover advantage, began to dominate the industry until early this century it completely took over. Typists preferred to learn on the keyboard that most employers used while employers bought the machines that most people were likely to know in order to minimise the cost of training. Therefore, a positive feedback system was set up where perfectly rational individual choices became translated into an irrational aggregate choice that belied the supposedly efficiency promoting free market system.

QWERTY still remains the standard keyboard layout even though in the 1930s it faced a new challenge from the Dvorak keyboard, which was promoted by the US Navy as being more efficient and capable of dramatically increasing typing speed; enough in fact to be able to quickly recoup the extra retraining costs. Nevertheless, Dvorak has been unable to budge QWERTY from its dominant position.

This example serves to illustrate the randomness inherent in the success of a technological design in a free market regardless of its technical merits or demerits. Indeed, in those technologies where the greatest cost is due to the investment of knowledge, there is likely to be increasing returns (positive feedback) driving the success of the technology once some initial advantage has been gained.

1.3 Exercises

1. (a) Discuss, with examples, which came first: Art, Technology or Science, and how they linked.
   (b) Would you classify engineering as a science or an art ? Why ?
   (c) Who do you think is responsible for bad technology: The engineer, the corporation that sells the technology, the politicians who allow it, or the consumer who supports it?
   (d) Has technology done more harm than good? Support your argument with examples.

2. (a) Make a count of all the technological products that you use from the time you get up in the morning to the time you go to bed (do not forget your bed !). Are you surprised at the number ?
   (b) Roughly how many of the products above were invented in the last 100 years? How many are dependent on electromagnetism ?
   (c) Which of the technologies can you absolutely not do without ? Why not ?
   (d) Which of the technologies can you live without (and maybe even wish were never invented)?

3. (a) What in your opinion has been the most important technological invention or development of the last 100 years? Of the last 1000 years ?
   (b) Give an example of an unintended consequence of technology and how that was resolved. Did the resolution itself have another unintended consequence ?

4. Read the summary of the article from Ref.[7] on the origin of the QWERTY keyboard for class discussion.
   (a) Did you know that a more efficient keyboard due to Dvorak was specifically designed to reduce finger movement, is faster to use and less tiring?
   (b) Compare the performance of the QWERTY and Dvorak keyboards at Ref.[8].
   (c) Did you know that you can easily reconfigure your keyboard to the Dvorak system? See Ref.[8].
   (d) Do you know why the letter ”R” is on the first row of the keyboard?
   (e) Can you give another example of technology lock-in ?

5. (a) Review the concepts of significant figures, scientific notation, and dimensional analysis. (See, for example the appendix of 2).
   (b) For a review of basic physics concepts see (3), and (4) or the “Hyperphysics” website (see link at course website) .
   (c) For soem other relevant figures for this module see (5) and (6).

1.4 References

1. Rhodes R, ed., Visions of Technology: A Century of Debate about Machines, Systems and The Human World, Simon and Schuster, c1999, ISBN: 0684839032

2. Giancoli D., Physics for Scientists and Engineers, Prentice Hall.

3. Gonick L . and Huffman, A., The Cartoon Guide to Physics, Harper-Perrenial 1991.

4. Hazen R.M. and Trefil J., The physical sciences : an integrated approach, Wiley , c1996.

5. Bloomfield L.M., How Things Work: The Physics of Everyday Life, Wiley 2001.

6. Wright M. and Patel M.N., eds. , Scientific American How Things Work Today.

7. Pool,R., Beyond Engineering, pp 156-161.

8. QWERTY at http://www.howstuffworks.com/question458.htm

Chapter 2

Newton's World

Any sufficiently advanced technology is indistinguishable from magic.
– Arthur C. Clarke

2.1 Newton's Laws

Newton, building on the observations and work of those before him, summarised his investigations of motion and change in terms of three laws:

1. **First Law: An object remains at rest, or continues its motion at constant velocity (i.e. constant speed in the same direction), unless acted on by an external force.**

   Speed is the rate of change of distance (x) with time, $s = \frac{\Delta x}{\Delta t}$, where we use Δ to represent a small change. If the object travels equal distances in equal times, then its speed is a constant and equals its average speed. However in general the speed of, for example, a car, changes during its motion and so we must consider its *instantaneous* speed at any particular time. This is obtained by calculating s in a small time interval Δt, and then taking the limit $\Delta t \to 0$. In that limit we use the calculus notation, $s = \frac{dx}{dt}$.

   Velocity defines not only speed but also the direction. Since direction can be specified by a vector $\mathbf{x}$ (also denoted as $\vec{x}$), velocity is also a vector quantity, given by $\vec{v} = \frac{d\vec{x}}{dt}$. So, for example, constant velocity means not only constant speed but also constant direction. A ball tied to the end of a string can be made to travel at constant speed around a circle but since its direction of motion is changing, its velocity is not constant.

   So what does the first law really say? It says that if you are in an *inertial frame*, that is one that is not moving or moving at constant

velocity, and if you observe an object changing its velocity, then some force is acting on that object.

The SI units of both speed and velocity are m/s.

2. **Second Law: The force acting on an object is equal to the product of its mass and acceleration, $\mathbf{F} = m\mathbf{a}$.**

What is acceleration ? It is the rate of change of velocity with time: $\vec{a} = \frac{d\vec{v}}{dt}$. So acceleration is also a vector, with units ms^{-2}. An object can be accelerating even if it moves at constant speed but changes direction. Notice that the force that must act on an object to cause it to accelerate is in the same direction as the acceleration.

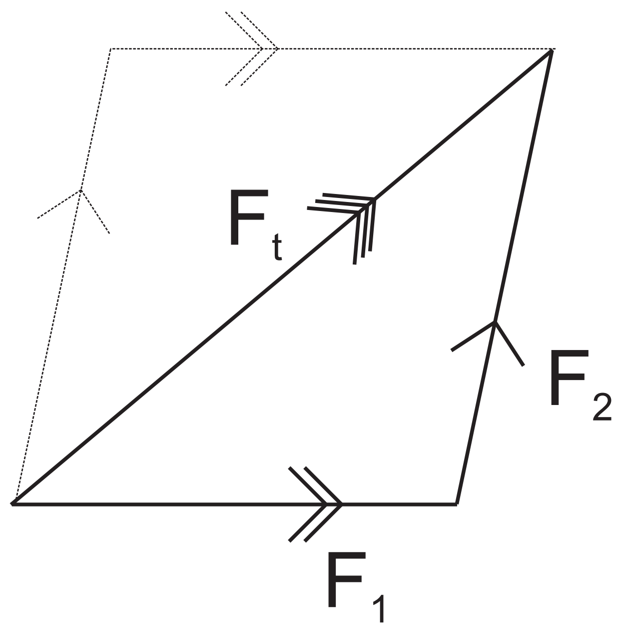


Figure 2.1: Addition of Forces

Example: In which direction is the force acting on the ball tied to the string mentioned earlier? Can you prove it ? What supplies the force ?

Newton's Second Law tells us how much force is needed to cause a certain acceleration or conversely what acceleration would result if a given force acted on an object. The parameter "m" appearing in the equation is the mass of the object, a measure of the objects *inertia* or resistance to change.

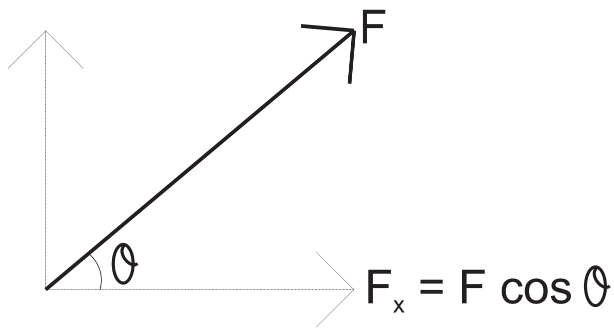


Figure 2.2: Resolution of Forces (i.e. Finding components along axis)

The SI unit of mass is the kilogramme, kg, and so the unit of force is $kgms^{-2}$. This basic unit is given the name **Newton**, N, in honour of Sir Isaac Newton.

3. **Third law: If one object exerts a force on a second object, then the second object exerts an equal but opposite force on the first object. (i.e. for every action there is an equal but opposite reaction)**

   Example: How does this law apply to the ball on a string example? Suppose you throw a softball with all your might, then why is it that you hardly move while the ball flies away? Does this contradict the third law?

2.2 Work and Energy

It is important to note that the word **work**, W, in physics has a very precise meaning different from its colloquial usage. The *work done on an object is defined as the force (F) acting on the object multiplied by the distance (d),* **in the direction of that force***, that the force acts on the body.* In symbols,

$$W = F \times d \tag{2.1}$$

The unit of work is therefore the "Newton-metre", which is also called the **Joule**, J, in honour of the physicist James Prescott Joule, a pioneer in the study of energy conservation.

Work performed on a body can be used to increase its *energy*. How much work is required to lift an object of mass m, a height h? If g is the acceleration due to gravity, then we know from Newton's Second Law that the force of gravity on the object is mg. Therefore we have to supply an equal but opposite force to move the object upwards a height h, doing an

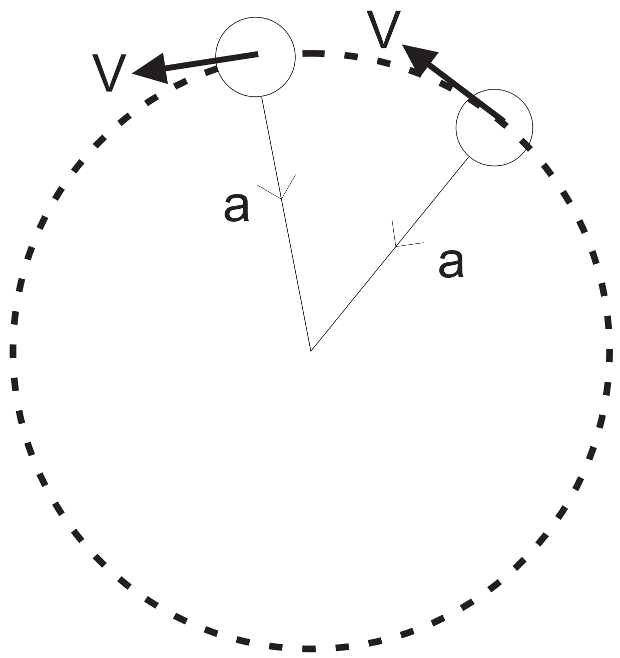


Figure 2.3: Velocity and Acceleration in Uniform Circular Motion

amount of work equal to mgh. This work, by definition, has been used to increase the potential energy of the object by mgh. That is,

$$U = mgh\,. \tag{2.2}$$

Energy due to position or the internal configuration of a material body is called potential energy, V. For example, a stone on a mountain top has gravitational potential energy due to its position, because when released it will roll down the mountain gaining speed, and hence kinetic energy, or energy due to motion. What is happening of course is that gravitational force does **work** on the stone, and this work causes the increase in kinetic energy of the stone. Conversely, to raise a stationary stone from the bottom of the mountain to the top requires us to perform work against the gravitational force. Then when the stone has been brought to the top, we say it has gained potential energy—the potential to increase its kinetic energy by rolling down the hill again.

Energy due to motion in space is called kinetic energy, T. Although we will not derive it in this course, it is useful to note that the formula for the kinetic energy of an object of mass m moving at speed v is given by

$$T = \frac{1}{2}mv^2. \tag{2.3}$$

Notice that a massive object has more kinetic energy than a lighter one, even when they are both travelling at the same speed.

Often one is interested at the rate at which work is performed, or the rate at which energy flows in some situation. **Power** is the rate of doing work,

$$P = \frac{\Delta W}{\Delta t} \tag{2.4}$$

The metric unit of power is clearly J/s and this is called a **Watt**, after James Watt, the person who invented the term "horsepower" to denote the power of his steam engines.

Careful observations and investigations by many researchers in diverse fields and over many years revealed that the abstract quantity called *energy*, while it may change in form in various processes, nevertheless its total quantity remains constant in a closed system, that is, it is conserved. The **Law of Energy conservation** is one of the most important laws of nature. In other words, the conservation of energy implies that if one takes the numerical sum of all the forms of energy of a closed system, then that sum is constant over time, in spite of all the changes in the forms of energy that may be taking place.

The ultimate source of most energy on Earth is the Sun. The Sun in turn generates its energy by a process of nuclear fusion, converting mass into energy according to Einstein's formula $E = mc^2$. It is estimated that the Sun will run out of fuel in about 5 billion years.

Force or Energy ?
Some of you might be confused about the difference between Force and Energy and why sometimes one concept or the other is used for explaining some phenomenon. Force, as you know is defined by Newtons Laws. It is the "cause" behind "cause and effect". The effect being acceleration etc. So in principle that is all you need: Find the forces acting on a body to determine its motion. Or knowing the motion of the body, you can deduce the forces acting on it.
However it is useful to introduce some additional concepts to help us discuss complicated situations. One of this concepts is energy: It is an abstract quantity that can take many forms and the Law of Conservation fo energy guarantess that energy in a closed system cannot be created or destroyed but only changed from one form to another.
How is energy related to force? Well if a force acts on a body and moves it some distance, then work is done. Carrying a ball up a flight of steps would require us to do work against gravity. The work done increases the energy of the system: In this case the ball gets gravitational potential energy. When it is dropped from the building, the ball picks up kinetic energy while losing potential energy. If there is friction, it also generates thermal energy etc. So the discussion of what happens can be conveniently be phrased in terms of energy. We could discuss the whole motion in terms of forces: When released , the gravitational force acts on the ball causing it to accelerate, and so on. The two ways of looking at it are equivalent and a matter of convinience.
Depending on the situation, a discussion in terms of forces might be more convenient, or one in terms of energy. It is easier to talk in terms of the energy in a battery and the power transferred by the battery rather than to talk about the force on the electrons caused by the electric field (the latter discussion is also interesting !).
May the Force (or Energy) be with You !

2.3 Placing a Geostationary Satellite in Orbit

Geostationary satellites are very useful: They remain above the same spot on earth and so can be used to bounce communication signals to and from locations far beyond the horizon without the need for cables or land-based relay stations. The fixed location of the satellites means that TV satellite dishes can remain pointing at the same location. Geostationary staellites are also used for weather forecasting.

The first satellite placed in a geostationary orbit was Syncom 3 in 1964 by the US Department of Defence.

There are many interesting questions one can ask about such satellites,

but let us begin with the following: How do they remain at one spot? That is, is it obvious that such a scenario is even possible?

The question above can only be settled by using the laws of physics. Let us start heuristically using an argument of Newton. Suppose one threw a rock horizontally. Then we know that it will follow a parabolic path to the earth because of the attraction of gravity. If we threw it harder and from a bigger height it would travel much further before hitting the earth. Now, the earth is not flat but roughly spherical, which means that if the rock is thrown hard enough horizontally from some height, then it a point would be reached where it dropped down at the same rate that the Earth curved away, so that it would keep going around the Earth!

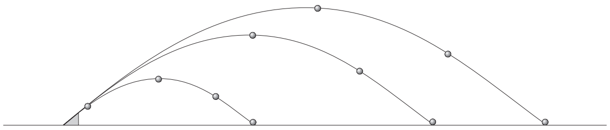


Figure 2.4: Motion of projectile relative to flat ground

Figure 2.5: Motion of projectile relative to curved surface

That is, just like a ball on a string can be made to go around by the tension in the string pulling it inwards, gravity pulling the rock towards the center of the Earth actually makes it go round the Earth if the rock has sufficient speed. Thus we see that in principle one could have an orbitting

satellite. Actually we have had a natural satellite for billions of years : The Moon! However, as we know, the Moon does not stay over one spot of the Earth. The reason being that the orbital period of the Moon about the Earth, about 28 days, does not coincide with the rotation period of the Earth which is one day (24 Hrs).

Hence, the question is, can we place an artificial satellite with an orbit of one day? Let us do the calculation. An object moving at fixed speed v around a circle of radius r has an acceleration towards the center, called the centrepetal acceleration, given by

$$a = v^2/r \,. \tag{2.5}$$

Thus by Newton's second law, a force of magnitude mv^2/r, where m is the mass of the satellite and v its speed, is required to keep the object in that motion. For a satellite that force is provided by the gravitational attraction between the Earth and the satellite, which is given by Newton's universal law of gravitation

$$\vec{F} = -\frac{GMm}{r^2}\,\hat{r} \,. \tag{2.6}$$

where $G = 6.67 \times 10^{-11} Nm^2/kg^2$ is the gravitational constant and $\hat{r}$ a unit vector in the radial direction pointing from the center of the Earth.(The law (2.6) holds for any particles of mass M and m separated by a distance r.)

The constant speed v is related to the orbital period T by $v = 2\pi r/T$. (The angular velocity ω is given by $\omega = 2\pi/T$, so that $v = r\omega$ and $a = r\omega^2$).

Matching the magnitude of the required force with that provided by gravity gives us

$$r = \left(\frac{GMT^2}{4\pi^2}\right)^{\frac{1}{3}} \tag{2.7}$$

as the distance from the Earth's center that the satellite must be placed to achieve a geostationary orbit. The distance comes to about 35,800 km above the Earth's surface!

(By the way, Eq.(2.7) is just Kepler's third law of planetary motion, which says that the cube of the distance is proportional to the square of the period).

Note that we knew qualitatively that placing an satellite in orbit was possible, what we needed the calculation for was to determine whether geostationary orbits *above* the earth were possible — if the Earth were much less denser for example there might not have been a useful solution (why ?).

2.4 Rockets

It requires a large amount of energy to place a satellite into orbit. Most satellites are launched using multistage rockets. The first stage consists of

engines that provide the huge thrust to lift the rocket and its load off the Earth and into flight. When the engines of the first stage have used up their fuel, the first stage of the rocket separates and falls to Earth while the second stage ignites and continues the process of placing the satellite into orbit.

For satellites to be placed in geostationary orbit, a third stage provides the final thrust to place the satellite at its final altitude after which another engine fires to give the satellite the required horizontal velocity to achieve a circular orbit.

Rockets work according to the principle of momentum conservation. Recall that the momentum of an object is the product of its mass and its velocity, $m\vec{v}$. Momentum conservation says that the total momentum of an isolated system is conserved. (If the system is not isolated, then the total momentum can change, and this change is affected by the external forces.)

Conservation of Linear Momentum

Newton's second law can be written in the following form :

$$\vec{F} = \frac{d}{dt}(m\vec{v}) \tag{2.8}$$

with $\vec{v}$ the velocity of the particle. The quantity $p = m\vec{v}$ is called the **momentum** of the particle, and so Newton's law states that the force acting on a particle is equal to its rate of change of its momentum.

Let us now consider a system of particles, with the particles labelled by an index i. Then Newton's law for the i-th particle reads

$$\vec{F_i} = \frac{d}{dt}(m_i\vec{v_i}), \tag{2.9}$$

and the sum (summation being denoted by the symbol $\sum$) over all the particles reads

$$\sum_i \vec{F_i} = \sum_i \frac{d}{dt}(m_i\vec{v_i}). \tag{2.10}$$

If the system is closed (isolated), meaning that no external forces are involved, then the sum of the forces muct vanish as according to Newton's third law for every action there is an equal but opposite reaction. Therefore we conclude that the total momentum of a closed system is conserved :

$$\sum_i (m_i\vec{v_i}) = \text{constant}. \tag{2.11}$$

The principle of conservation of linear momentum says that if there are no external forces acting on a system, then regardless of the complicated interactions among the particles, their net momentum will remain constant. This principle explains rocket propulsion and other situations discussed in the exercises.

Rockets burn an internal supply of fuel which is then expelled through a nozzle at the back: The rearward momentum of the exhaust gases must be balanced by an equal forward momentum change which is what propels the rocket forward. It is important to note that the rocket does not need to push against any material object (such as air, which is absent in outer space) to move forward.

The nozzle expels the gases in a particular direction and the rocket moves in the opposite direction. The nozzle used on large rockets has an interesting shape: The reason for the flaring is that it permits the exiting gas to expand and thus reduce its pressure, allowing a greater conversion of random thermal energy into the directed kinetic energy of the exhaust, and hence a larger forward thrust.

In order for a rocket to overcome the forces of gravity and escape Earth's orbit, it must attain a sufficiently high velocity: Recall that the higher an object is above the Earth, the greater its potential energy, hence the rocket has to convert chemical energy to kinetic energy which is then traded for the increase in potential energy. Those high kinetic energies are possible only through large accelerations caused by the ejected burnt gases. It is usual and economical to use multi-stage rockets whereby each stage is discarded as its fuel is used up.

A rocket engine is the most powerful, compared to propeller power and jet-engines, for its weight. In addition to large space rockets used for launching satelites, there are also uses for smaller rockets such as for powering pilot ejection seats, missiles and fireworks. You can read more about rockets in Ref.[**??**].

2.5 More on satellites

2.5.1 Power Supply

A satellite must obviously carry its own power source to carry out its mission which may last many years. Solar cells are most commonly used. The solar cells are usually mounted on winglike structures that open up from the body of the satellite once it is in position. The solar cells of a Global Positioning System satellite occupy about 5 sq m in area and generate about 700 watts of electricity. Batteries are also carried to provide power when sunlight is not available, and during the initial phase before the solar panels are operational.

The solar panels of course should be pointed towards the Sun while the antennas and sensors of communications satellites must be pointed earthward. The orientation and stability can be maintained by various means. One is the use of gyroscopes (See for example Chap.3 of [1]), large spinning wheels which maintain angular momentum and resist changes in direction.

2.5.2 Survival under Extreme Conditions

When a satellite faces the sun it experiences intense heat and alternatively intense cold when it is hidden from the Sun. In addition the equipment on the satellite itself generates heat which must be dissipated to avoid damage to the electronics. In outer space the only way this heat can be removed is by radiation and so satellites usually have adjustable panels to allow the heat to radiate.

The satellites on board computer must also be protected from damaging cosmic radiation. However over the long run the satellite does suffer damage from cosmic radiation and dust (tiny meteroids).

2.5.3 Did you know?

There are about 2500 functional satellites orbitting the Earth, and if one counts also all the space debris such as discarded rocket stages the number is about 8000!

The first satellite launched into space was the Sputnik 1 in October 1957. It weighed just $83kg$ and stayed in orbit for 95 minutes. A month later Sputnik two became the first vehicle to launch a living creature, Laika the dog, into orbit. Unfortunately Laika was sacrificed in the name of science. The USA managed to catch up only in 1958 with a satellite launch of its own. The space race betwen the Soviet Union and USA was on with the former leading in most new landmarks, such as the first man to orbit space, the first man to walk in space and the first probe to the moon. In 1961 John F. Kennedy committed the USA to a Moon landing and massive resouces were channelled for that purpose. Apollo 11 landed on the moon on July 20 1969 and the first moonwalk by Neil Armstrong took place. The manned lunar programme came to an end in 1972 with Apollo 17.

The space programme cost billions of dollars but the push to achieve the goal resulted in many new technologies being developed which were later commercialised, that is, useful spin-offs resulted. You can read more about these on the Internet.

2.6 Centrifuges

There are many technologies that make use of centripetal acceleration. These may be generally called centrifuges ("centrifugal force" is the fictitious force in the rotating frame of reference that appears to push objects away from the center). Examples are merry-go-rounds, the loop-the-loop of of roller coasters, spin-dryers and centrifuges in laboratories.

The centrifuges used in laboratories are used to separate components of different sizes in a fluid mixture. Ultracentrifuges can generate accelerations up to 100,000 g's, i.e. 100,000 times the acceleration due to gravity.

Conservation of Angular Momentum
Another very useful conservation principle is that of **angular momentum conservation**: In an isolated system, the total angular momentum of the system is conserved. Angular momentum is associated with rotational motion, and for a mass m rotating at a distance r from an axis at an angular velocity ω, the angular momentum is given by $L = mr^2\omega$. For extended objects, the formula is $L = I\omega$, where I is called the moment of inertia of the object and is simply the summation of mr^2 over the different elements making up the object.
The law of conservation of angular momentum is the reason why an ice-skater rotates faster when she pulls in her arms: Notice how a remarkable and useful conclusion can be drawn even for a complicated system as an ice-skater.
Just as linear momentum is the tendency for an object to keep moving at constant speed in a straight line unless acted on by an external force, angular momentum is the tendency for objects to keep rotating at a constant angular speed unless acted on by an external *torque*, which is just the external force multiplied by the perpendicular distance from the axis of rotation.
Just like linear momentum, angular momentum is a vector and its conservation implies both direction and magnitude.

2.7 Summary

Here is a summary of some topics from this chapter.

- Newton's Three Laws
  First Law: An object remains at rest or moves at constant velocity unless acted on by an external force
  Second Law: The force acting on an object is the product of its mass and its acceleration.
  Third Law: If an object exerts a force on another object, then the second object exerts an equal force in the opposite direction on the first object.

- Work and Energy
  Work done: Work done on an object is the result of a force acting on it to move it a certain distance. It is the product of the force and the distance the object is moved *in the direction of the force*.

  Potential Energy: Energy inherent in a body due to its position or internal configuration. For example a compressed spring or an object

in a gravitational field.

Kinetic Energy: Energy of an object due to its (bulk) motion.

Power: Rate at which Work is done.

Conservation of Energy: The total energy in a closed system can be neither destroyed nor created but only transformed from one form to another

- Circular Motion and Satellite Orbits
  An object moving at constant speed around a circle experiences a centripetal acceleration towards the centre. therefore a force is required to keep an object moving in circular motion at constant speed!

  Newton's Law of Gravitation states that the force of attraction between two point masses is directly proportional to their masses and inversely proportional to the square of the distance between them.

  A geostationary satellite remains fixed relative to one point on Earth's surface. GPS satellites are not geostationary.

- Some Conservation laws
  Conservation of Linear Momentum: The total momentum of an isolated system (i.e. if no external forces act) is conserved. This is the principle by which rockets move in outer space.

  Conservation of Angular Momentum: The total angular momentum of an isolated system (i.e. if no external torques act) is conserved. This is the principle by which a ballerina spins faster when drawing in her arms. It is also why tail rotors are needed on helicopters.

2.8 Exercises

1. An Ox, which has heard of Newton's third law (but not understood it very well), refuses to pull a cart attached to it because it argues that : "If I pull on the cart with a certain force, Newton's third law would imply an equal but opposite force, which would then mean that the total force on the cart is zero, and so it will not move. So why waste my effort? I won't pull". What's wrong with the Ox's argument?

2. (a) A car is moving at constant speed along a straight road. Draw a diagram to indicate (direction) all the forces acting on the car and explain the origin of those forces. (NOTE: do not confuse forces acting ON the car with forces actiong on the road etc. So you are drawing what is called a "free body diagram" for the car).
(b) Is the friction between the tires good or bad?
(c) What determines the force of friction between the road and the tires?

3. Describe two ways in which you could make yourself seem (from the scale reading) lighter when standing on a weighing scale. How would you make yourself seem heavier ?

4. You are sitting in a car which is moving smoothly at constant velocity.
(a) If you held your book at head level and released it, what would the path it took to your lap look like from your perspective, and also to someone stationary on the road?
(b) If you threw your book vertically up, along which trajectory would it fall back ?
(c) What forces act on the book in the above experiments?
(d) What would happen in part (a) if the car accelerated ?

5. You are sitting in a car which makes a sharp right turn.
(a) In which direction would you tend to move ? Would you say that a force was acting on you ?
(b) How would the situation look to someone stationary observing you from the road?
(c) How is the analysis in part (a) and (b) consistent with Newton's laws?
(d) Why is it that the top part of your body moves more than the bottom ?

6. A ping-pong ball is released from a height of 3m onto a hard surface.
(a) What height will the ball rebound to and why ?
(b) Describe all the energy conversions that take place from the moment the ball is released to the time it finally comes to rest.
(c) Repeat the above exercises for a ball released on the Moon.

7. (a) What is the kinetic energy of a sprinter of mass 70 kg running at 10 m/s ?
(b) What is the kinetic energy of a car of mass 1000 kg travelling at 60 km/hr ? (c) The car in (b) is brought to a stop by applying a constant force over a distance of 10 m. Determine the magnitude of the force.

8. A car traveling at high-speed requires a large force to stop it.
(a) How is it possible for unassisted (non-power) brakes to apply such a large force?
(b) Is energy conserved? Where does it go?
(c) Which moves a greater distance: The foot-pedal or the brake-pads? Why?

9. (a) Why do cars skid when making a fast turn on the road ?
(b) How do racing cars avoid that problem ?

10. When you fire a rifle, you feel recoil force. How is this related to the principle of momentum conservation ?

11. You, of mass $60kg$, are rollerblading along the East Coast at $10m/s$ when you notice two kids in front of you blocking your path with their bicycles. You manage to come to a stop in 0.7 seconds to avoid a collision.
(a) What was your initial and final momentum ?
(b) Why is momentum not conserved ?
(c) Where is the origin of the relevant external forces (if any) ?
(e) What is the magnitude of the force that was required to stop you ?
(d) In practise how would you come to a stop ?

12. Which is easier to stop with your bare hands and why: A bicycle, or a car, both travelling at $1m/s$.

13. Why do bumper cars in amusement parks have thick bumpers ?

14. Two ice-skaters are intially at rest. The larger one, of mass 70kg pushes off against the other of mass $50kg$.
(a) What is the total momentum of the skaters before and immediately after the push ?
(b) If the larger skater moves off at a speed of 3 m/s, what is its momentum ?
(c) Determine the speed of the smaller skater after the push-off.

15. Where is the "crumple zone" in a car loacted, what role does it play, and what are the principles on which it operates?

16. List and briefly discuss the technologies in modern cars that reduce the possibility or extent of human injuries in the vent of a collision.

17. If an astronaut exits an orbiting space-shuttle to carry out repairs, does he have to worry about falling to earth or moving away from the shuttle?

18. (a) Using the known values of G, $M = 6 \times 10^{24} kg$ and T, calculate r in equation (2.7).
(b) Using the known radius of the Earth (6377 km), determine the height of a geostationary satellite above the equator.
(c) Determine the orbital speed of the satellite.
(d) In which direction, east to west or vice versa, does a geostationary satellite have its orbital direction ?
(e) In which plane does can a geostationary satellite be placed ? (i.e., e.g., can it be placed outside the equatorial plane ?)
(f) How do countries not along the equator make use of geostationary satellites?
(g) The Space Shuttle orbits at about $200km$ above the Earth's surface to avoid the atmosphere. Determine its orbital period.
(h)Optional : How might you determine the values of G, M and the Earth's radius ?

19. (a) Why is it that you sometimes feel "light" or "heavy" while in a fast moving lift?
(b) Why is it that you momentarily feel "light" when the car you are in (or the skateboard you are on) speeds over a large bump in the road ?

20. (a) If a centrifuge of arm-length 0.5m produces up to 50,000g's, how fast is it rotating ?
(b) Juice of mass 0.01kg is placed in a blender. If it is made to rotate at a speed of 0.5m/s along a circle of radius $r = 0.1m$, then what inward force must the sides of the blender exert on the juice? How does the force compare with the weight of the juice?

21. (a) Why does water leave the clothes in a spin-dryer rather than accelerate with the clothes?
(b) When the water escapes the centrifuge, in which direction relative to the centrifuge does it travel ? Why ?

22. (a) Small rockets can be used in space to change the direction of a satellite or other craft. Since there is nothing to push against in empty space, how do the rockets accomplich their task ?
(b) An initially stationary rocket of mass M ejects a small amount m of matter (hot gas) at high speed V. What will the final speed of the rocket be ?
(c) During one blast, the rocket expels 20 kg of gas out of the back with an average velocity of 100 m/s. What is the change in momentum of the rocket ?

23. (a) Is the space-shuttle a single-stage or a multi-stage vehicle ? Where if any, are its multiple stages ?
(b) Man went to the moon using large multistage rockets, but the return trip needed only a small rocket. Explain why.

24. (a) How does an ice-skater begin her spinning motion ?
(b) What would be the isolated system (consisting of the skater) to which you could apply the law of conservation of momentum ?

25. Optional: Here are some general knowledge questions.
(a) Do you know who had the idea for communication using geostationary satellites? See, for example, Ref.[2].
(b) Which was the first satellite launched and how big was it ? What was it intended for and what did it achieve? See Ref.[3].
(c) Which was the first living creature sent to space by Humans?
(d) What are the different purposes to which satellites have been used ? (See, e.g., Ref.[**??**]).
(e) What would be the most significant change in your life if satellites never existed ?

2.9 References

1. HTW wiley website at : http://htw.wiley.com/htw/
2. Clarke biography at http://www.lsi.usp.br/∼rbianchi/clarke/ACC.Biography.html
3. Sputnik at http://www.hq.nasa.gov/office/pao/History/sputnik/

Chapter 3

Fluids: Balloons, Pumps and Submarines

Unlike solids, fluids such as air and water do not have a fixed shape but adapt to the shape of the container or surroundings they are in. Many technologies depend on their function on the special properties of fluids. In this chapter we will study those properties characteristic of static fluids and the way they constrain technologies.

3.1 Buoyancy and the Archimedes Principle

Consider some fluid, such as water in a large container, in equilibrium. As in the figure, consider now an imaginary volume marked out to define an object, and let us look at the forces acting on that volume. There is the downward force due to gravity equal to the weight of the volume, which must be balanced by a net upward force due to the surrounding fluid since we have an equilibrium situation. Thus the upthrust, or buoyancy force is equal to the weight of the volume.

What would happen if instead of the imaginary marked volume we had an actual object of that shape immersed in the fluid? Since the upthrust is a property of the fluid and not the object, it will be the same and equal to the weight of the volume displaced! *Thus the upward buoyancy force acting on an immersed (partially or fully) object is equal to the weight of the fluid displaced.* This is known as **Archimedes Principle**.

If ρ_0 is the density of the fluid, V_0 the displaced volume and g the acceleration due to gravity, the upthrust is equal to $\rho_0 V_0 g$. On the other hand, the downward force on the object is equal to its weight, which is $W = \rho V g$, where V is the volume of the object and ρ its density. Since an object will float only if the upthrust can balance the weight, this can only happen if $\rho = \rho_0 V_0 / V \leq \rho_0$.

That is, for an object to float in water its density must be less than

that of the water. This is how pieces of wood, or rafts, boats, canoes, and metal ships are able to float, with the part of them under water displacing just enough fluid to give the upthrust to balance their weight . Note that if the density of the object becomes equal to the density of the fluid then the object can be fully submerged and remain at equilibrium. This is how ballons and submarines operate !

3.2 Balloons

Let us look at balloons. Since in natural circumstances they are fully immersed in the air, we conclude that whether they will sink, remain floating, or move upwards, depends on whether the ballons density is greater than, equal to, or less than, that of the surrounding air.

If the balloon is filled with air, then it will sink in air because its average density includes the density of the balloon's material (say rubber for small ones, silk or nylon for larger ones). Thus the ballon must be filled with a gas such as Helium which is less dense than air. The density of air near the surface of the Earth is about $1.25kg/m^3$ while that of Helium is 0.14 of that value. Thus a Helium filled balloon will move upwards when released.

But even theoretically, the balloon cannot keep going up forever. The reason is that density of air decreases as we go upwards, as those of you who have gone mountain climbing know. Why is this so? Well, just think of our atmosphere: It is mainly concentrated in a belt of 6km in thickness, held there by the force of gravity. But why doesn't gravity pull all the air down to the ground? Actually it tries to, but as more of the air is pulled down, the pressure and density at ground level builds up which creates an opposing upwards force. Thus in equilibrium in the atmosphere reaches a density profile which is approximately exponentially decreasing with height.

Thus the buoyancy force on a balloon decreases with height. For hot-air balloons as discussed below, this sets a limit to how high a baloon can go. For gas-filled balloons, such as those used to gather wind speed and direction data, the skin will rupture when it has expanded sufficiently at large heights.

The pressure due to air at ground level is about $100,000N/m^2$ (or "Pa" after Pascal). This large pressure acts on our bodies but we do not realise it because our bodies produce balancing internal forces. Only when there is an imbalance, such as when we climb a mountain or are in a plane, that we notice some discomfort.

Hydrogen as a balloon gas is cheaper and more abundant than Helium, but is also flammable and so can be dangerous. Note that though hydrogen is half as dense as helium, its lifting capacity is only slightly better than that of helium.

3.2.1 Hot-Air Balloons

Helium is relatively expensive and not practical for large balloons like those used to carry adventurers. Indeed the solution that predates the use of Helium is hot-air. We all know that hot air is less dense than cold air (Do you ? What evidence do you have?) and so can be used to lift the balloon and its occupants. The first hot-air balloon carrying human passengers was built by the Montgolfier brothers in 1783. Such balloons were also used for observation by armies in World War I. The first round-the-world balloon trip took place in 1999 lasting about 19 days, using a hybrid helium-hot-air balloon.(See the Figure).

Let us look at the situation more quantitatively. Recall that real gases at moderate pressures and temperatures obey universal gas laws, the Boyle's law and the Gay-Lussac-Charles law, which can be summarised as saying that $PV = a(\theta + \theta_0)$ where a and θ_0 are constants and θ is the temperature in the Celcius scale where $0°$ is the freezing point of water and $100°$ is the boiling point. An extrapolation of the empirical curve to the limit $PV = 0$, which is allowable for an ideal gas, gives the intercept $\theta_0 = 273.15$ and so it is convenient to define an absolute temperature scale T, measured in Kelvins K, by

$$T = \theta + 273.15\,. \tag{3.1}$$

to simplify the ideal gas law to

$$PV = NkT\,, \tag{3.2}$$

where P is the pressure, V the volume, $k = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant and N the number of atoms/molecules in the gas. The coldest temperature one can attain is $0K$, **absolute zero**, where all molecular motion would cease.

The air inside a hot-air balloon is heated by an external burner located near the open end of the balloon. The ideal gas law shows that even though the pressure of the air inside and outside the balloon may be balanced, which in fact allows the end to remain open without losing much of the hot air, the partcle density $n = N/V$ is less for the hot air than for cold air. That is, less molecules of hot air are needed compared to cold air to achieve the same pressure. Thus the density of hot air is less than that of cold air, allowing the ballon to float.

The buoyancy of the hot-air balloon can be controlled by controlling the burner and thus the temperature of the air inside. Notice how the buoyancy problem is solved differently for the hot-air and helium balloons.

Airships evolved from balloons and had large envelopes enough to provide sufficient lift to carry many passengers, plus engines and propellers for propulsion. Such airships are now used for advertising or some millitary early warning and surveillance purposes.

3.3 Water Pressure

Unlike air, water is an incompressible fluid. This means that pressure applied to water at one part of an enclosed system gets transmitted to other parts. The incompressible nature of water also makes the calculations of the effects of water pressure easier. For example, the pressure due to a column of water of height h is clearly given by

$$P = \rho g h \tag{3.3}$$

where ρ is the density of water and g the acceleration due to gravity. Since the density of water is $1000kg/m^3$ and $g = 9.8m/s^2$, the pressure due to a $10m$ water column is about $100,000$ Pa which is normal atmospheric pressure. That means that normal atmospheric pressure (due to kilometres of air!) can support a column of air 10m high in a vertical sealed tube that is evacuated of air above. Thus in water the pressure increases by about $10,000$ Pa for every $1m$ of depth. Hence $10m$ below sea-level the pressure on an object is about 2 atmospheres. These facts also imply that pumps must be placed at the bottom of a well hole if it is required to raise water more than $10m$ high.

Water in a network of tubes will flow to equalise the height of water since the pressure depends on the height. The dependence of pressure on height also complicates the distribution of water to very tall buildings. The water must be pumped up against gravity to water tanks at the top, but then it cannot be supplied to all floors directly from the same pipe as the top floors might have very low water pressure while the lower floors have dangerously high levels. Pumps and water towers help to maintain high-pressure while regulators help to dissipate the energy of fast moving water.

3.4 Water Pumps

Water pumps exert pressure on water to make it flow in a pipe or from a region of low pressure to one of higher pressure, as from a reservoir to the distribution network. The figure shows a simple reciprocating pump.

The human heart has two pumps that are resposnsible for blood flow between the lungs and the heart and the rest of the body.

3.5 Compressible fluids

One advantage the compressibility of gases has is the ability to store energy as thermal energy and to release that energy in expansion by doing work, as we shall see in the working of engines. Aerosol cans work on the same principle as do shaken soda/beer cans.

3.6 Exercises

1. In the last chapter we learnt that a satellite could be placed in geostationary orbit above the Earth only at a certain height.
   (a) Why is it then that in the absence of wind, balloons can stay at stationary points above the earth at much lower altitudes?
   (b) Why is it that for the satellite problem we could ignore some of the physics issues that crop up for the balloon?
   (c) Why not use balloons as communication satellites?
   (d) Other than recreation, are there any useful uses of balloons?
   (e) If a helium ballon was released on the Moon, where would it go ?

2. There are plans to build a self-contained floating city, with permanent residents, that will encircle the globe every two years. The "Freedom Ship" (Ref.[1]) will be about 1.3 km long, 200m wide and 100m tall. It will weigh about 2.7 million tons.
   (a) How much water will the ship displace?
   (b) How much of the ship will be under water?

3. (a) Waterproof watches have a maximum depth to which they can safely be taken while swimming. Why?
   (b) Why must tall dams be much thicker at their bases than their tops?

4. (a) What causes water to rise up a drinking straw?
   (b) Why can't you drink from a long straw if you're more than 10 m above the drink?
   (c) How do vacuum cleaners create the "vacuum and how does this allow them to draw in matter?

5. Water can be drained from a container using a siphon.
   (a) How does a siphon work?
   (b) The toilet tank uses a siphon, valve and float to operate and regulate the flush. Explain how this is done. (Hint: See Ref.[3])

6. Suppose that the water pressure at a fire-hydrant at street level is 500,000 Pa above atmospheric pressure. Firemen have connected a hose to the hydrant to use in battling a fire in a tall building.
   (a) How high can the firemen take the hose inside the building and still expect water to flow out of it?
   (b) If instead, the firemen stay at street level and spray the water up, how high will the water go?
   (c) One way to boost the water pressure is to send it through pumps in the fire-engines. If a single fire-engine is used to boost the pressure by 500,000 Pa, what is the result for the above two questions?

7. Describe the action of piston pumps such as those used in water pistols and bicycle pumps. (Hint: See p132 of Ref. [4].)

8. A car traveling at high-speed requires a large force to stop it.
(a) How is it possible for unassisted (non-power) brakes to apply such a large force?
(b) Is energy conserved? Where does it go?
(c) Which moves a greater distance: The foot-pedal or the brake-pads? Why?

9. (a) Why are there water tanks at the top of tall buildings?
(b) At what time of the day would the pumps be working to pump the water up? Why?
(c) Would it be desirable for a skyscraper to have one tank at the top or several at different levels? Why?

10. Read about "Submarines" in Ref.[5] and then answer the following questions briefly.
(a) How do submarines change their buoyancy under normal conditions to dive or surface?
(b) What constrains the depth to which submarines can dive?
(c) How far down can modern submarines go?
(d) What pressure does water exert on a submarine submerged to the depth you answered in above?
(e) How does a submerged submarine navigate under water?
(f)Where does a submarine get its supply of fresh-water for long journeys?
(g) How is the air inside a submarine kept breathable?
(h) How is the moisture that is exhaled by the crew removed? Why must it be removed?

3.7 References

1. Freedom Ship at http://www.howstuffworks.com/floating-city.htm

2. Lienhard J., The Engines of Our Ingenuity, 262p.

3. Toilets at http://www.howstuffworks.com/toilet.htm and "531/2 Things That Changed The World", by D.West and S.parker pp13.

4. Bloomfield L, How Things Work: the Physics of Everyday Life, Wiley 2001 2nd Ed, ISBN: 0471381519

5. Submarines at http://www.howstuffworks.com/submarine.htm

Chapter 4

Motion Through Fluids

Heavier-than-air flying machines are impossible.
—- *Lord Kelvin*

Objects can be "suspended" above the ground in at least two ways: They can be in free fall towards the ground which is moving away just as fast (satellites in geosyncrhonous orbits), or they can be held in static equilibrium by the buoyant force of fluid (balloons). Here we will consider a third way in which dynamic equilibrium can be achieved: Using Bernoulli's principle. (**Caution**: Fluid dynamics is a very complicated subject, and the explanations below are highly simplified and for the most part qualitatively correct. More detailed and accurate explanations exist but are beyond the scope of this course. see e.g. Ref.[**??**])

4.1 Bernoulli's Equation

As pumps exert a force on the water in a pipe, they do work which is carried by the water. For an *incompressible fluid in steady state flow* (which we assume for now on), part of that energy is carried as kinetic energy of the water. If the water is flowing on a level pipe of varying diameter, the kinetic energy of a unit volume increases where the pipe is narrower as it must flow faster there, and conversely the kinetic energy decreases where it slows down at wider regions. Since energy must be conserved, what is changing in opposition to the kinetic energy is the *pressure potential energy* of the water along the level pipe. Pressure potential energy is the product of pressure times volume. (We have also assumed that there are no viscous forces, so no energy is lost as heat).

If in addition the level of the pipe changes, then the gravitational potential energy also changes and must be taken into account in the energy balance equation. The net result is Bernoulli's Equation:

$$p + \rho g h + \frac{1}{2}\rho v^2 = \text{constant along a streamline}, \tag{4.1}$$

where ρ is the density of the fluid, g the acceleration due to gravity, h the height difference, v the velocity, and p the static pressure. A streamline is the path that a particular volume of water takes in steady state flow.

What Bernoulli's equation says, for example, is that one can exchange kinetic energy for gravitational potential energy (say in a water fountain) or exchange gravitational potential energy for pressure (say in water towers).

Consider now streamline flow of an incompressible fluid in a pipe which decreases in diameter in the direction of flow. For steady state flow, the volume of liquid passing through any cross-section of the tube in one second must be the same, and hence the liquid must flow faster through the narrower region. Bernoulli's equation then implies that the static pressure of the fluid in the narrower region must be less than in the broader region.

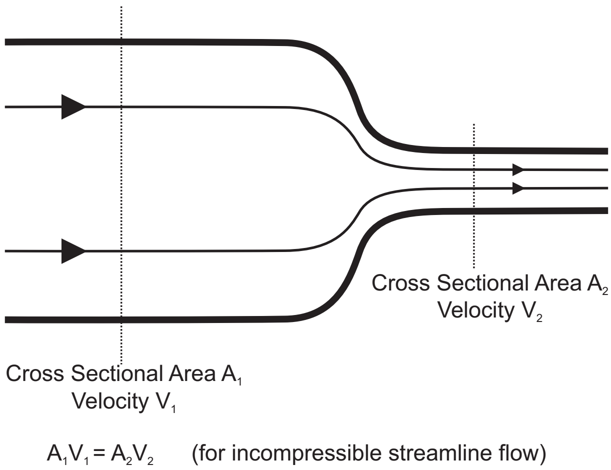


Figure 4.1: Streamline flow of an incompressible fluid

4.2 Airplanes

Bernoulli's equation implies that, for streamline flow, the pressure drops when a fluid speeds up. Instead of moving fluid, one can consider an object moving through the fluid, such as an airplane wing shown in the figure.

In the frame (that is, from the perspective) of the wing, the air is moving pass it. The air must deviate and take a longer route over the top convex surface compared to the lower surface. That is (ignoring the slight compressibility of air here), the streamlines over the convex portion are closer together, representing faster fluid flow compared to the bottom surface.

This means that the pressure will be less on the top surface compared to the bottom surface of the wing, resulting in an upward lift force.

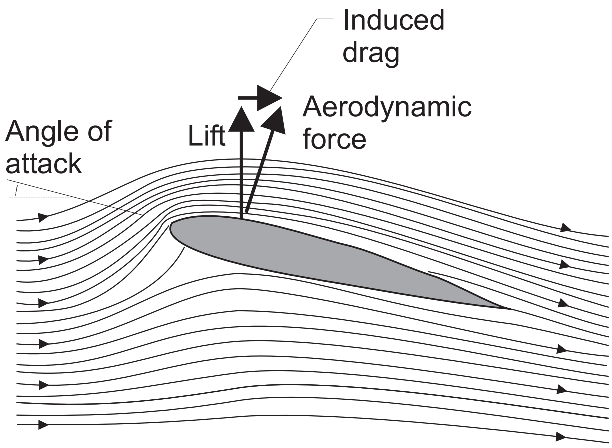


Figure 4.2: Forces acting on an airplane wing

Thus a plane, which is heavier than air, remains afloat not by the buoyant force of air, but by an aerodynamic lift force.

Actually, the net force on the wing is not vertically upwards but rather has a horizontal component against the direction of motion of the wing. This **induced drag** tends to slow the plane down. One way to understand the origin of this induced drag is as follows: The upward lift on the plane implies an equal but opposite downward force on the air. Since the air was originally flowing horizontally, its downward deflection requires energy that must be supplied by the plane. Hence the plane must lose some energy in the process and this is identified with the retarding force.

In order to minimise the energy loss associated with induced drag but still obtain sufficient lift, one may try larger wings which move more air but deflect it slowly so that they carry away less kinetic energy. However one must note that larger wings imply a greater **viscous drag** associated with friction between the air and the surface of the wing.

The lift on the wing increases with the airspeed and the curvature of the top surface (See,e.g. [5]. As we shall see later, jet planes are able to fly faster than propeller driven planes, and hence the former can make do with smaller and slightly curved wings. Actually, commercial planes have adjustable parts to their wings, called slats and flaps, which can be extended during take-off to increase the area and curvature fo the wings. Consequently the plane does not require too high a speed (or an excessively long runway) to generate sufficient lift.

There is another trick to obtaining greater lift with a low airspeed: If the plane is tilted so that the wings make a larger angle ("angle of attack") with

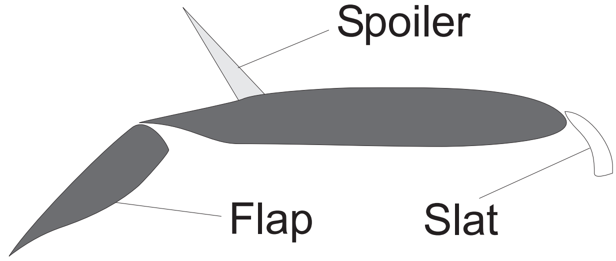


Figure 4.3: Wing sections

the airstream, then the greater is the differential in airflow on the upper and lower surfaces, creating a greater lift. However there is a limit to how much lift one can obtain simply by increasing the angle of attack of a slow moving aircraft: At some point the airstream forming the **boundary layer** of the wing might detach from the wing, leading to a turbulent region above the wing and to a consequent rise in air-pressure there. The wing therefore loses its lift power causing the plane to drop.

The ailerons on a plane's wing tips, and the elevators and rudders on the tail are used to steer the plane by deflecting air flow. The ailerons are used to rotate the plane about its fuselage axis, the elevators to raise or lower the nose, and the rudder to make it turn left or right.

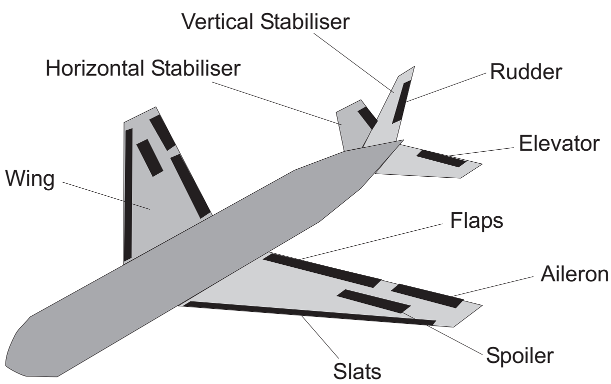


Figure 4.4: Parts used to steer a plane

Man has always been fascinated by flight, trying through the ages to emulate the birds. Though many people had build contraptions to try and achieve this dream, it was the Wright brothers who went about the quest in a systematic manner, combining scientific theory with experimental testing, to develop a controllable and stable flying machine. You can read a summary of their procedure in Ref.[5].

Wing-In-Ground Effect
During the Second World War pilots who had a damaged engine would fly lower, just a little over the sea, to save on power. By flying close to the ground, greater lift is acheived due to the high pressure air cuhion trapped between the wings and the ground.You may have noticed the effect, called the “wing-in-ground effect” (WIG), in movies where small planes sometimes seem to “hang” in the air as they are about to land (or when a sheet of paper seems to slow its descent just before it hits the ground). During the Cold War the Russians conducted a top secret research programme to build massive planes that would fly close to the ground to save on fuel and possibly also try to evade radar. Recently Boeing has shown interest in building similar large planes for potential commercial and millitary applications. The Japanese are experimenting with using the WIG effect for fast trains, hoping to achieve greater efficiency than magnetically levitated trains.

4.3 Propellers

So a plane can obtain lift if it is moving through air using large curved wings. But how does a plane move through the air in the first place? One way is to use propellers.

As can be seen from the figure, the shape of propellers is such that when they rotate, the configuration is exactly like that of plane wings moving through the air. Instead of a lift force, now the propellers provide the necessary thrust to move the plane forward. Not surprisingly, propellers have many features in common with the wings discussed earlier, namely the thrust increases with size, speed of rotation, curvature and angle of attack.

Propellers have several limitations and problems. Just as for wings, they suffer induced drag which therefore necessitates a continuous input of energy from the engine turning the propeller. A second peculiar problem is that while initially the propeller generates thrust in stationary air, as the plane moves it must continue to generate thrust in air which is moving towards the propeller from the front. The solution is to have the propeller blade swivel forward to meet the incoming air at the right angle. The blade itself is often shaped to enable its operation under varying air velocities.

Human Powered Flight
On August 23 1977, The Gossamer Condor became the first human powered aircraft to complete a prescribed figure 8 course around two turning pts 800m apart and win a prize of 50,000 pounds. The designer Dr. Paul MacCready then built the first human powered aircraft to cross the English Channel and also the first solar powered aircraft. He has also designed some increbible hand-held toy gliders: You can watch the video "Flying Free" on these latter inventions.

4.4 Viscosity, Boundary Layers and Turbulence

When discussing the flow of fluid in a pipe and Bernoulli's equation, we ignored the viscoity of the fluid and assumed the flow to be streamlined or "laminar", in contrast to turbulent flow whereby initially nearby regions of fluid frequently collide and move far apart. In a real fluid, various factors determine whether one has laminar or turbulent flow.

Consider first the case of **internal flows**, such as fluid flow in a pipe. A dimensionless quantity, called the **Reynolds number** can be defined as:

$$\text{Reynold Number} = \frac{(\text{density of fluid}) \times (\text{diameter of pipe}) \times (\text{speed of flow})}{\text{viscosity of fluid}} \tag{4.2}$$

Low Reynolds number corresponds to laminar flow: In this regime the viscosity (internal friction) of the fluid keeps it together. By contrast at high Reynolds number the inertia (density) of the fluid dominates motion and it becomes difficult to maintain order between neighbouring elements of the fluid, resulting in turbulent flow. For Reynold numbers below a critical threshold of of 2000 $\sim$ 2300, the fluid flow remains laminar with minor pertubances being smoothened out. However above the critical value, it is difficult to maintain laminar flow and turbulence generally develops.

For **external flows**, such as fluid motion over a ball's surface or over the surface of a wing, one can also define a Reynold's number but the critical values depend very much on the shape of the body.

4.4.1 Sphere

Consider the case of a small smooth ball thrown slowly. For low Reynold's number ($\sim< 4$) the motion of air is laminar and symmetrical around the ball as shown in the figure. In particular the front and back of the ball are regions of low air speed and high pressure while the sides are region of high air speed and low pressure. The symmetrical situation means that the ball does not feel any **pressure drag**. However this does not mean there is no

resistance to the motion of the ball through the air. Just as we discussed for the case of airplane wings, there is still viscous drag on the ball due to friction between the ball's surface and the air at the boundary.

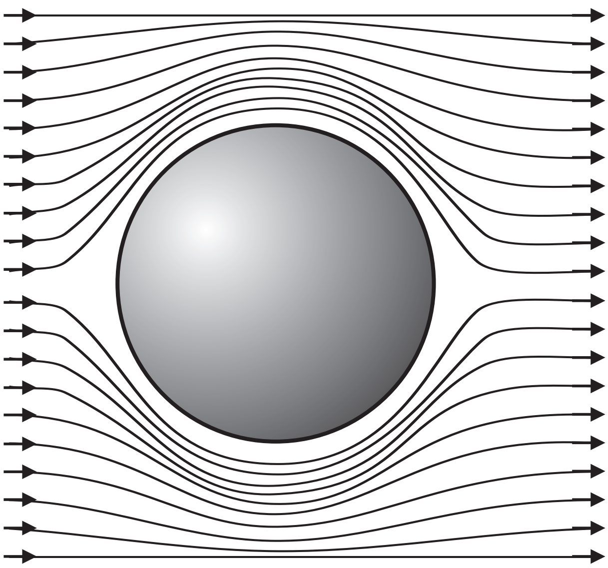


Figure 4.5: Streamline flow around a slow moving ball

Indeed the viscous forces between the surface of the ball and the surrounding air layer create a small **boundary layer** of air that is moving much slower than the air further away, as shown in the figure. For low Reynold's number the entire airstream is able to flow smoothly from the front to the back of the ball because the viscous forces between the air layers are strong enough to pull the boundary layer with them to the back where (as we have noted) the pressure is higher than the sides of the ball.

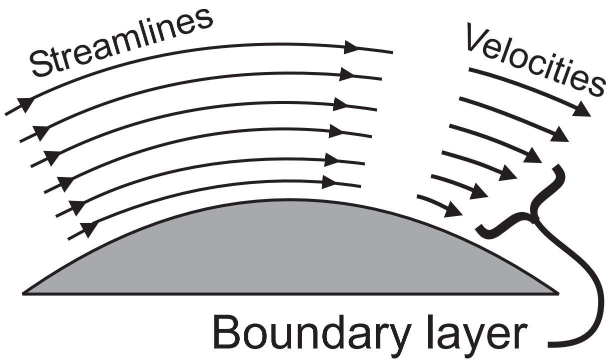


Figure 4.6: Boundary Layer

However as the Reynold number increases, a point is reached whereby the viscous forces between air layers are insufficient to pull the slower moving boundary layer with them. Consequently the boundary layer does not quite make it to the back of the ball, colliding with and disordering the air behind the ball. The chaotic motion of the air behind the ball disrupts the perfect symmetry that held for laminar flow. In fact now the pressure behind the ball becomes lower than in the front. This region of low air-pressure is called an air-pocket or turbulent wake. *The net effect is that the ball experiences pressure drag and can not travel as far.* For $4 < Re < 40$(*which practically means any ball used in sports*), the airflow becomes mildly turbulent while the turbulence gets more pronounced for $Re > 130$.

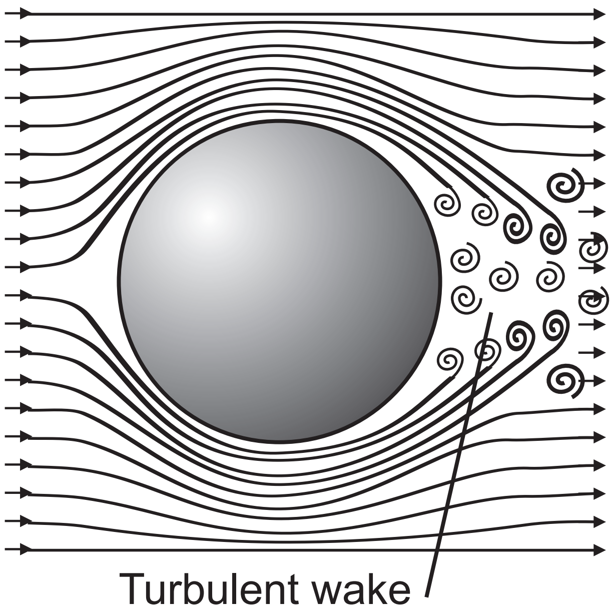


Figure 4.7: Turbulent wake leading to pressure drag

Amusingly, the solution to the problem of pressure drag is to create much more turbulence! The idea is to make the boundary layer turbulent so that it can mix with and get energy from the air-layers further away. This will then allow the energised boundary layer to travel further towards the back of the ball, reducing the turbulent wake and thus the pressure drag on the ball. For a smooth ball the boundary layer becomes turbulent for $Re \sim 5 \times 10^5$. In practise the boundary layer can be triggered into a turbulent state at much lower fluid speeds: That is why tennis balls have fuzz on them.

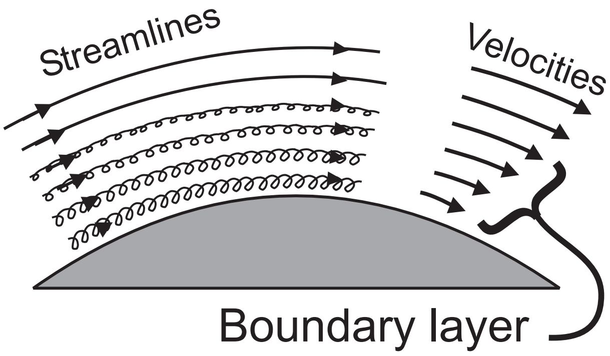


Figure 4.8: Turbulent boundary layer

4.4.2 Cylinder

The flow around a smooth and long cylinder is similar to that of a sphere but with some interesting differences because of the different symmetries of the objects. Eddies form when Re exceeds 4 but the wake is still laminar until $Re > 40$ when the wake becomes unstable and two staggered rows of vortices form, moving downstream with opposite sense of rotation. The pattern is called "Karman vortex street". For $Re \sim 80 - 200$ the eddies attached to the cylinder break off periodically and alternately from the two sides of the cylinder resulting in a lateral vibration of the cylinder. Such vibrations can be dangerous.

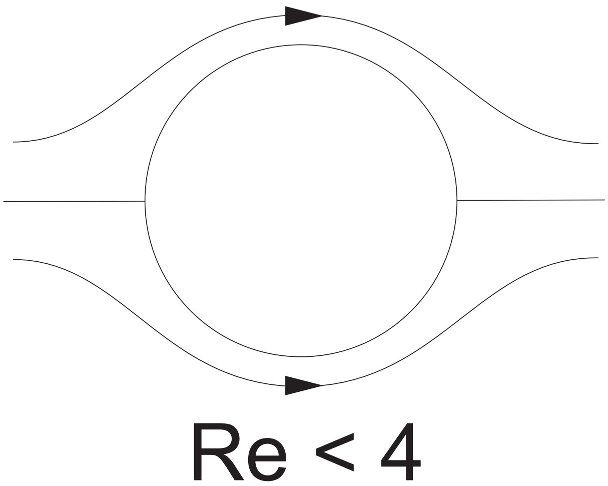


Figure 4.9: Streamline flow around a cylinder

Just as for the sphere, for a high value fo the Reynolds number the pressure drag actually decreases as shown in the figure.

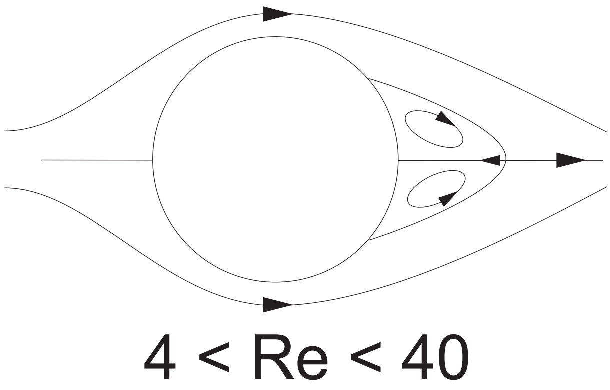


Figure 4.10: Formation of Eddies

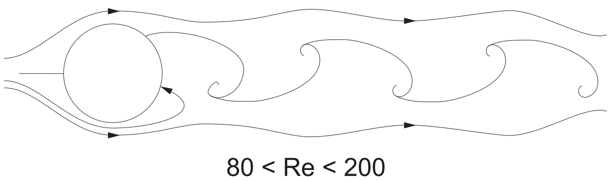


Figure 4.11: Karman Street

4.5 Curve Balls

In many sports balls are made to curve by giving them a spin. A full discussion of the forces acting on such balls is quite complicated because of turbulence which generically develops, so here we will simplify the problem by assuming that the air flow is still laminar.

Suppose you throw a ball forward and give it a spin which is clockwise when observed from the top. Then, since the ball carries a boundary layer of air with it, the left side of the ball will have lower net velocity compared to the right-hand side. The air-pressure on the left-hand side will therefore be higher than the the air-pressure on the right-hand-side and the ball be be deflected to the right.

One of the most remarkable cases of a curve ball occured in a soccer match in 1997 between Brazil and France. Roberto Carlos kicked the ball with such force and spin that it went far right of the wall of defenders and seemed to miss the goal completely. However the ball curved sharply near the end and entered the goal! See Ref.[3] for more on the physics of curve balls.

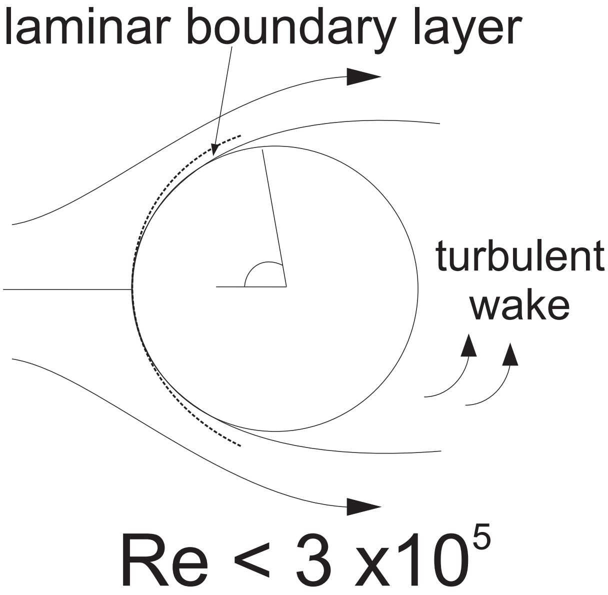


Figure 4.12: Turbulent wake behind cylinder

4.6 Drag

We have mentioned various forms of drag, or resistance to motion, that an object can experience in a fluid. Let us summarise them here:

- Viscous (or friction) drag, caused by friction between the fluid and the surface of the object.

- Pressure drag, caused by the formation of a turbulent wake behind the object and the consequent pressure drop at the back relative to the front.

- Induced drag, which arises for assymmetrical objects whereby the fluid gets a net deflection from its original direction of flow.

The figure shows how the percentage that viscous drag and pressure drag make up of the total drag on some bodies. Notice how streamlining a shape can reduce pressure drag but increase viscous drag.

In general it is neither possible nor convenient to consider the three forms of drag separately and one usually just defines a drag coefficient for a particular shape. The drag F_D (force opposing motion) is related to the drag coefficient, C_D, by the relation :

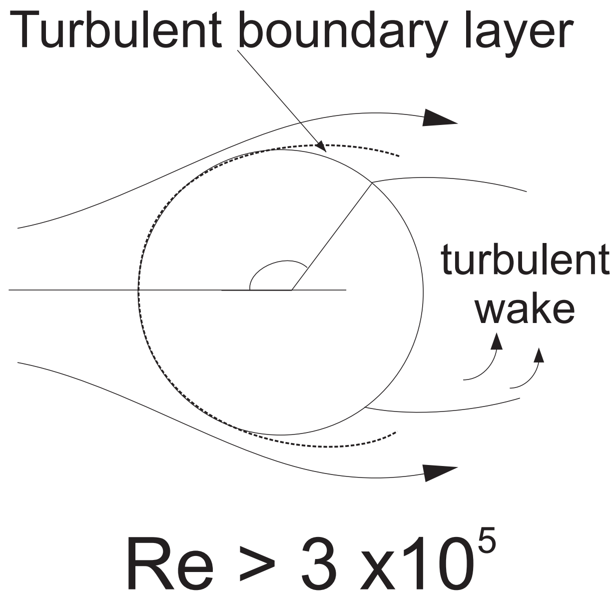


Figure 4.13: Turbulent boundary layer around cylinder

$$F_D = C_D\ A\ \frac{\rho U^2}{2}\,, \tag{4.3}$$

where A is the projected area of the body on a plane perpendicular to the direction of fluid flow, ρ is the density of fluid and U is the flow velocity. As mentioned above, the drag coefficient depends on the shape considered and also on the speed (or Reynolds number).

Similarly, for an airfoil, one has the equation for the lift (upward force),

$$F_L = C_L\ A\ \frac{\rho U^2}{2}\,, \tag{4.4}$$

where A is area of the wings, ρ is the density of fluid and U is the flow velocity. The lift coefficient, C_L depends on the shape of the wing, the speed (or Reynolds number) and the angle of attack.

The drag and lift coefficients are determined experimentally and tabulated or plotted as graphs for use.

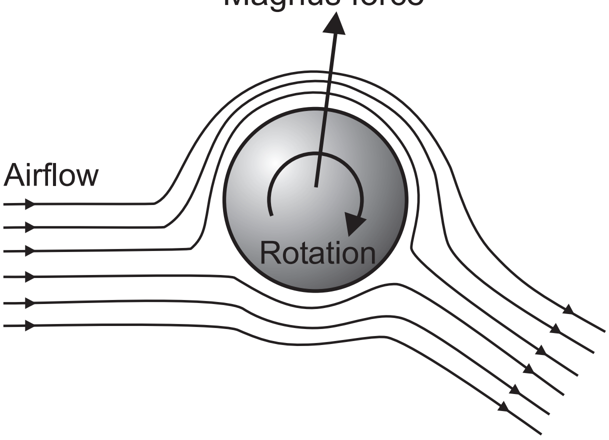


Figure 4.14: Forces acting on a spinning ball

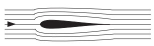
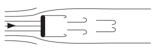
Shape	Pressure drag D_p (%)	Friction drag D_f (%)
	0	100
	≈ 10	≈ 90
	≈ 90	≈ 10
	100	0

Figure 4.15: Pressure and viscous drag for various shapes

Once again: Force or Energy ?
As for how air is deflected by an airfoil and how lift is created, one can use Bernoullis principle which is really a statement of the conservation of energy for (for approximately incompressible and non-viscous) fluid flow. Or one can directly use Newton's laws and look at the force the air exerts on the upper and lower surfaces of the foil etc. Which one is more convenient depends on the situation: Both must give the same right answer if used correctly.
Similarly drag or the resistance an object experiences due to its flow through a fluid can be sometimes conveniently discussed in terms of forces acting on the body (e.g surface friction = viscous drag, or pressure differential at front and back = pressure drag) or in terms of energy expenditure (e.g. deflecting air downwards =¿ giving up some energy =¿ foward motion kinetic energy reduced (unless engine works harder) =¿ equivalent to having a force against motion = induced drag).

4.7 Jet Engines

Larger thrusts and greater speeds can be attained by using jet engines instead of propellers. The engines on large commercial planes are called turbofan engines, which are turbojet engines to which a giant fan is attached to the front.

Let us therefore first describe the workings of a turbojet engine shown in the figure. Air which enters the inlet is compressed by a series of fans to many times atmospheric pressure before it is mixed with fuel and ignited. The resulting hot mixture has even higher pressure now and tries to expand, rushing out of the nozzle at the back of the engine. Since on the whole the incoming air has been accelerated out of the engine, by Newton's third Law, there is an equal but opposite force on the engine pushing it and the plane forward.

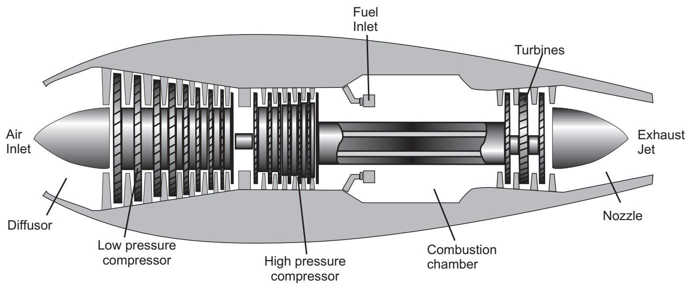


Figure 4.16: Turbojet engine

As the burnt fuel mixture is rushing out of the back of the engine, it is made to pass through a fanlike device called a turbine which drives the compressors at the front of the engine.

The second figure shows a turbofan engine. The big fan in front of (and powered by) the turbojet engine compresses large amounts of incoming air, most of which is then expelled through the ducts at higher pressure thus propelling the plane forward. Some of the air is fed into the turbojet engine and is used as described before.

For some historical background on jet engines see Ref.[4-6].

4.8 Summary

- Assuming an incompressible fluid in steady state flow (no turbulence), and assuming no viscous losses, the conservation of energy of the fluid is given by Bernoulli's Equation:

$$p + \rho g h + \frac{1}{2}\rho v^2 = \text{constant along a streamline}, \qquad (4.5)$$

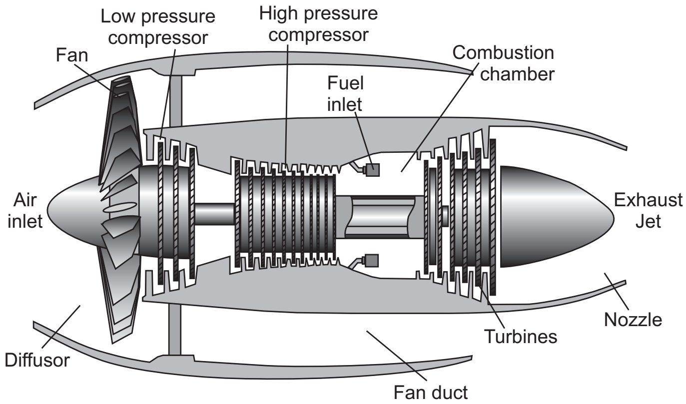


Figure 4.17: Turbofan engine

- Airplanes
  1. An aerofoil generates lift because the air flowing over the top will be moving faster and hence have a lower air pressure than the air stream flowing along the bottom surface.
  2. The amount of lift can be increased or decreased by increasing or decreasing the angle (called the angle of attack) at which the aerofoil meets the oncoming air flow.
  3. Ailerons are flaps on the wings that rotate the plane about its length, elevators are flaps on tail that raise or lower the nose of the plane and the rudder is a flap on the tail fin that turns the plane right or left.
  4. A propeller is an arrangement of multiple aerofoils that constitute the propeller blades but unlike the wings meets the oncoming airflow head on to generate forward thrust.

- The Reynolds number is a dimensionless quantity that measures the likelihood of a laminar (straight) fluid flow becoming turbulent.

4.9 Exercises

1. (a) Jet planes on aircraft carriers have to take off and land on relatively short runways. Various innovations such as catapults, braking cables, and slanted decks were introduced to achieve this. Explain how and why those devices work.
(b) Harrier jump jets are another solution to the problem of short runways. Discuss how they function.
(c) Can you think of other solutions to the above-mentioned problem ?

2. (a) How does a pilot change the direction (left, right, up ,down) that a plane is heading?
(b) What changes happen to the wings of a commercial plane as it takes off ? (c) How does a landing plane reduce its speed ?
(d) What is the role of the tail and fins on a plane ?
(e)What is the role of the feathers on a shuttle cock, and the fins on a dart or arrow ?

3. (a) Commercial planes flying high above the Earth need to pressurise the air inside the plane so that the passengers can breate normally. Where does the compressed air come from and how is it compressed?
(b) A pilot can determine his speed of travel using a pitot tube, shown in the figure. Explain how it might work.

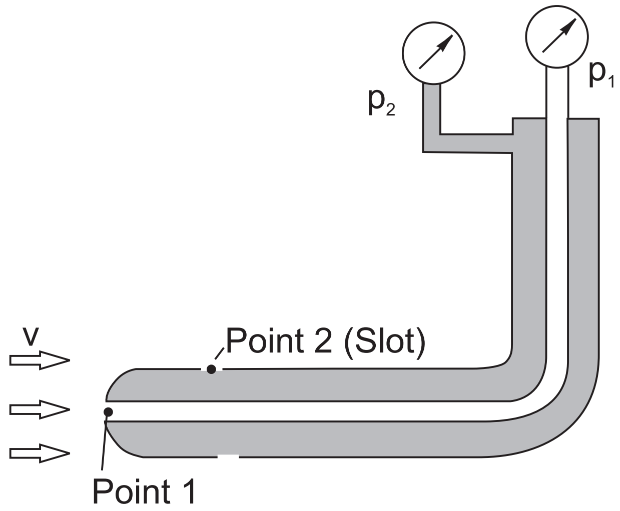


Figure 4.18: Pitot tube

4. Gliders are able to travel large distances without any engines.
(a) Why is an engine not required to keep a glider afloat ?
(b) What limits the distance a glider can travel ?
(c)Optional: Why do paper airplanes look different from real ones ? (See Ref.[2])

5. (a) What factors determine the amount of lift an airplane wing generates?
(b) Experiment with the free NASA aerodynamics software on the web and compare qualitatively with the formula in the text.

6. Look carefully at the shape and orientation of the blades of a ceiling fan, stand-fan or table-fan in your home. Observe also the direction

of rotation of the blades when the fan is switched on.
(a) Hence explain how and why the proper functioning of the fan depends on the correlation of the blade shape/orientation and direction of rotation of the fan.
(b) What is the difference between the fan and a propeller on an aircraft ?
(c) Is there a force on the fan ? Why and in which direction ?

7. (a) Explain how the propellers on a helicopter enable it to generate lift.
(b) What is the function of the small tail roor on many helicopters?
(c) Some helicopters do not have a tail rotor. How do they solve their problem then ?
(d) How does a helicopter produce thrust to move forward ?
(e) What are the advantages and disadvantages of a helicopter compared to an airplane?

8. (a) Explain how the propellers on a boat or ship function to propel the craft.
(b) Can boats use a fan above the water instead of a propeller underwater to propel them forward ? When would this design be useful ?

9. It has been stated that a great deal of damage to houses exposed to hurricanes and tornadoes is due to forces from *within* the house pushing outwards. Does this make sense? Explain.

10. Two possible spins that might be given to a ball are topspin or backspin.
(a) Describe the aerodynamic deflection the ball might experience in each case as a result of those spins.
(b) When might you use those spins ?

11. (a) Read the article on aerodynamics design of cars for class discussion.
(b) Racing cars have some features borrwed from airplanes. What are they and how do they help the racing cars?

12. (a) Regarding the imperfections introduced into golf and tennis balls, can you think of other ways of introducing the imperfections to reduce pressure drag?
(b) In particular, discuss why tennis balls do not have dimples.

13. Why do you feel stronger wind some distance above the ground than close to the ground ?

14. One place where turbulence is useful is when fluids have to be mixed. Explain why turbulence is helpful in this case and how it can be achieved.

15. (a) Estimate the maximum speed at which a smooth (dimple-lesss) golf-ball can travel through the air without creating turbulence. Is this a natural speed for golf-balls?
(b) At what high speed must the smooth golf-ball travel to take advantage of the dramatic reduction in pressure drag ? Do you think this speed is attainable in practise? (Justify your answer)
(c) Repeat part (a) and (b) for movement of a smooth ball through water. (Some numbers : Density of air is about 1.25 kgm^{-3} and its viscosity is $1.8 \times 10^{-5} kgm^{-1}s^{-1}$. The viscosity of water is about $10^{-3} kgm^{-1}s^{-1}$.)

16. Why does turbulent flow of water, such as that coming out of a water hose, create noise ?

17. Why is it that cyclists in competition try to ride behind one other as much as possible ?

18. Planes flying close to the ground to exploit the WIG effect seem to have many benefits. What drawbacks have delayed their commercial implementation ?

19. Not all inventions are designed intentionally. Many useful artifacts have evolved from an initial accident or spin-off from different applications. Optional: Read the examples in Ref.[7] to educate and entertain yourself.

20. Consider the following statement: *In general, older technologies tend to serve recreational or aesthetic roles once the replacement (by newer technologies) is complete. This was the case with fuel wood, sailing ships, horses, convertible (open) cars, and many other examples." — N. Nakicenovic (1986).*
(a) Give another example in support of the statement.
(b) Think of a counterexample. What causes are responsible for your counterexample?
(c) Are there examples of older technologies co-existing with newer technologies that perform similar functions? How did this come about?

21. Optional : Read pp 28-42 of Ref.[5] and Chapter 7 of Ref.[6], on the systematic design of the original airplanes by the Wright Brothers and how modern airplanes are designed using computers.

4.10 References

1. How Things Work, by L. Bloomfield
2. Paper airplanes at http://www.howstuffworks.com/airplane8.htm
3. Physics of curve balls http://physicsweb.org/article/world/11/6/8
4. "531/2 Things That Changed The World", by D.West and S.parker pp44-45.
5. Pages 28-42 of: Voland, G,. Engineering Design, Addison-Wesley, 1999.
6. Petroski,H., Invention by Design.
7. Flatow,I., They all Laughed: Velcro: pp 131-134, Artificial Sweeteners: 156-160, Video Games: pp213-221.
   Jones, C.F. Accidents May Happen: Kites pp18-19, Yo-yo pp24.
   Jones, C.F. Mistakes That worked: Frisbee pp 35-36, Piggy Bank, pp36-37.

Chapter 5

Thermodynamics and The Environment

Murphy's Law of Thermodynamics:
Things get worse under pressure.

In this chapter we look at engines that obtain their energy from burning fuel and other artifacts of industrialisation that consume energy, and what their impact is on the environment.

5.1 The Internal Combustion Engine

By contrast to external combustion engines, like the steam engine, which burn their fuel on the outside and then use the steam to move parts inside the engine, an internal combustion engine such as that used in a car burns its fuel inside the engine itself.

A typical internal combustion engine operates on a four-step (Otto) cycle:

- Induction Stroke
  The piston moves down the cylinder creating a partial vacuum and the inlet valve opens to introduce air and and some fuel, such as petrol. A car engine will have at least four of these cylinders which act to convert the chemical energy of the fuel into thermal energy and then mechanical work.

- Compression Stroke
  During this step, the inlet valves close while the piston moves forward compressing the fuel-air mixture. Once maximally compressed, the mixture is ignited by the spark plug.

- Power Stroke
  The burning mixture reaches high-pressure and expands to push the piston back and perform work on the car.

- Exhaust
  The burnt and relatively cool remains are squeezed out as the outlet valve opens and the piston pushes inward again.

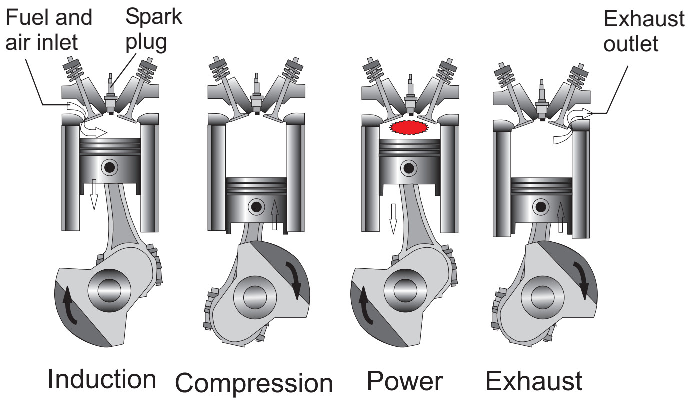


Figure 5.1: Four-stroke internal combustion engine

The cycle repeats, with the linear motion of the piston converted by the crankshaft into rotational motion that drives the wheels.

Of course one has to first start the engine somehow. This is done by the starter motor which is powered by a battery.

How efficient is an internal combustion engine ? That is, what fraction of the input energy from the fuel can be converted into useful work? Obviously in any real engine there are losses due to such factors as friction in the moving parts, leakage of heat, and incomplete burning of the fuel. However it turns out that even a perfect engine would have an efficiency of less than one! In order to understand why this is so, we have to review some basic concepts of thermodynamics.

5.2 Heat is a form of Energy

In Chapter 2 we had learnt that in a closed system, energy is conserved. That is, one can transform energy from one form to another such as potential energy to kinetic energy, and use such energies to perform mechanical work (recall the physics definition of work).

In the nineteenth century, many physicists believed that heat was a form invisible fluid, called "caloric", stored in any substance. This hypothesis seemed to explain the observed flow of heat from one body to another and

the fact that some bodies seem to have more heat than others. However the caloric theory makes other predictions which can be tested. For example, it implies that the amount of heat (the caloric fluid) in any substance is a constant. On the contrary, Benjamin Thompson (also known as Count Rumford), found that the amount of heat generated when boring canons was not fixed but depended on the friction between the moving parts: That is, the heat generated depended on the amount of work performed. Thus the caloric theory had to be abandoned.

The work of Thompson and others culminated in the classic experiment of Joule, who demonstrated in a set-up how falling weights converted their potential energy to kinetic energy of a wheel, which turned a paddle in a water. The water heated up as a result, leading Joule to establish the equivalence of heat to mechanical energy.

Thus heat is just another form of energy, and in a heat engine such as the internal combustion engine, the objective is to transform this heat energy into mechanical work. Actually, heat is also a form of motion : The motion of **atoms and molecules** that make up a substance. The atoms that make up matter are always jiggling about to some extent and their mean kinetic energy is a measure of the amount of thermal energy in the object.

We know from experience that different objects can have different amounts of thermal energy as seen from the heat “flow” from a “hot” object to a “colder” object. Indeed **temperature** is used to quantify the direction of heat flow between two objects in contact. If there is no net heat flow, one says the objects are in thermal equilibrium, while heat flow when it occurs is always from the object at higher temperature (hot) to one at lower temperature (cold).

For scientific purposes it is convenient to use the absolute temperature scale, T, measured in in **Kelvins**, K. The conversion between the temperature T and the familiar Celcius scale ($\theta°C$) is given by

$$T = \theta + 273.15\,. \tag{5.1}$$

5.3 Laws of Thermodynamics

The study of thermal energy and its transformations led to the formulation of various laws of thermodynamics. These are

- Zeroth Law
  If two objects are each separately in thermal equilibrium with a third object then they are also in thermal equilibrium with each other.

  The Zeroth Law is used everytime we use a thermometer.

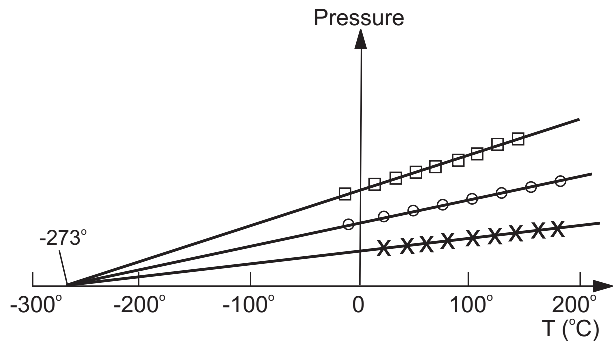


Figure 5.2: The Ideal Gas Law

- First Law
  The change in internal energy of an object is equal to the heat input minus the work done by the object.

  Thus the **First Law of Thermodynamics** is just our familiar Law of Conservation of Energy, which states that energy, of which heat is one form, cannot be created or destroyed but only transformed from one form to another. Heat energy if not converted into mechanical work, increases the objects **internal** potential energy.

  The conservation of energy disallows the existence of "perpertual motion machines of the first kind": Machines that can function perpetually in a cyclical fashion without any input of energy. So, for example, you cannot run your car-engine without fuel. (If a car running on a smooth road has its engine turned off, it can still continue some distance until losses due to friction with the road and air-resistance consume its energy and it comes to a stop: Without friction or resistance the most the car can do without any energy input is to continue in a straight line.)

  However not all conversions which are allowed in principle by the First law are possible in practise! For example we know from experience that it is impossible to construct an engine that would work cyclically by *only* taking in heat and transforming it *completely* into mechanical work. If such a thing were possible we, in our equatorial land, would have been able to design such an engine to extract heat from the hot air around us, cooling our surroundings in the process, while using the converted energy to surf the Internet — paradise!

  The impossibility mentioned above disallows "perpetual machines of the second kind", and is encoded in the Second Law of Thermodynamics, of which one version is :

- **Kelvin's Statement of the Second Law**
  It is impossible to construct a cyclical engine that solely transforms heat completely into work.

  Kelvin's statement can be shown to be equivalent to the following:

- **Clausius Statement of the Second Law**
  It is impossible for an engine to solely take heat from a colder reservoir and deliver it completely to a hotter reservoir. (That is, heat will not flow spontaneously from a cold body to a hot one).

5.3.1 Order and Disorder

In previous sections, we mentioned that heat energy is just the random kinetic motion of molecules, whose mean kinetic energy is a measure of their temperature. So Clausius statement just says that when two bodies are in contact, it is unlikely that the molecules in the hotter body will move even faster while colliding with the slower moving molecules of the colder body. In fact we naturally expect that due to the molecular collisons, the molecules in the colder body will speed up and those in the hotter body slow down, implying a heat transfer from the hot to cold body.

Clausius's statement sounds reasonable, but how does one understand the limitations placed by the Kelvin statement? Again, note that heat energy is a disordered form of energy compared to ordered forms such as gravitational potential energy or bulk kinetic motion. While it is easy to create disorder out of order, the reverse is very difficult: Because the number ($\sim 10^{23}$) of molecules in any macroscopic piece of matter is very large, it is statistically improbable for disordered molecules to spontaneously act in concert to create order. One says that heat energy is of lower grade than ordered forms of energy and the Second Law places limitations of the conversion of low grades of energy into higher grades.

To understand the the idea of order and disorder, consider the following example: An isolated box contains an equal number of two types of molecules, say 'white' and 'black'. Suppose one had the extreme case where initially all the white molecules were on the left and all the black molecules on the right. Of course this is an unnatural situation and soon, due to collisions, the molecules will totally mix. We say that the initial situation corresponds to a state of maximum *order* since there are fewer ways for the molecules to arrange themselves, while the final situation corresponds to large disorder since then there are many more ways for the molecules to arrange themselves(each molecule can now be anywhere in the box).

We also know that if the number of molecules is very large, it is extremely unlikely that the gases will revert to the totally separated state at some future time. Thus the increase in disorder in the system is for all practical purposes **irreversible.**

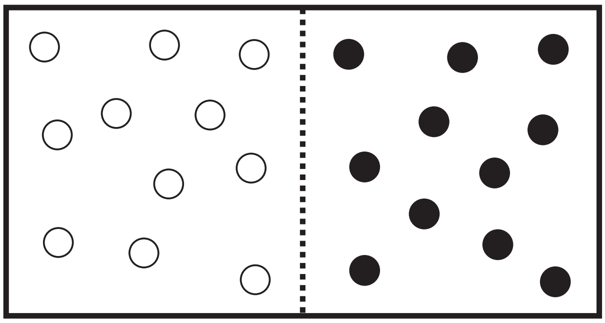


Figure 5.3: An unlikely ordered state

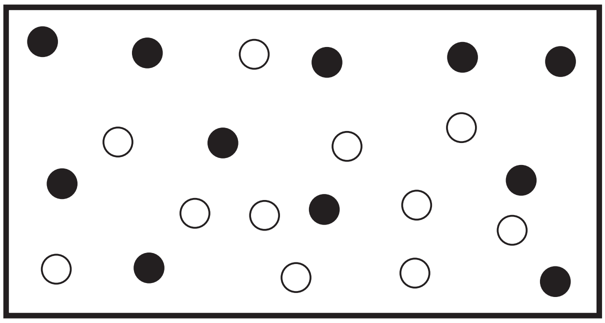


Figure 5.4: A likely disordered state

Note that the irreversibility is not due to the underlying fundamental laws (e.g. Newtons laws are reversible) but a result of the system going from an unlikely ordered state to a more probable disordered state, and the fact that for large systems (large number of molecules), the probability of the system reverting to the ordered state being negligible. In the above example, the probability of any one molecule being on the left half of the box is $1/2$. If there are N molecules, the probability that all of them are on the left is $(1/2)^N$. Even for N as small as 100 this works out to be about 10^{-30}, an infinitesimal quantity ! For macroscopic materials, N is of the order of 10^{23}, and so the resulting probability is even lower.

A quantitative measure of disorder in a system is given by the **entropy** of the system, with the entropy of a disordered state much higher than the entropy of an ordered state.

The above example and others like it lead one to the following equivalent form of the **Second Law of Thermodynamics**: *The entropy of a thermally* **isolated** *system never decreases.*

Thermally isolated systems are also called closed systems: There is no exchange of energy or matter with the environment.

As we have seen from the examples above, The Second Law defines the direction in which events in **isolated** systems can proceed: **Isolated systems tend to decay and become more disordered, increasing their entropy** (or keeping it constant but never decreasing it !). Here is another example: If sugar is placed in a cup of hot tea, it soon dissolves and one will not find the situation whereby the isolated system consisting of the tea and sugar spontaneously separates into its constituent parts. Similarly, if some ink is spilled in a glass water, it soon spreads and colours the whole glass.

It is very important to note, from the argument given in the last subsection, **that the Second Law is actually a statement about average behaviour that becomes overwhelmingly likely in a very large system, meaning that exceptions will be unobservable in all practical situations**.

Thus heat energy is of a higher entropy, or lower grade, than other forms of ordered energy like gravitational potential energy. The Kelvin statement places limitations on how much work a cyclical engine can extract out of disordered energy like heat – some of the input heat energy must be dumped into a lower temperature reservoir and is wasted. The limits placed by the Second Law are of practical concern and will be discussed below.

5.3.2 Entropy

One can derive the expression for entropy in thermodynamics but we simply state a result here without proof.

Suppose a very small quantity of heat ΔQ is *added* slowly to a system at fixed temperature T, then the entropy S of the system is *increased* by

$$\Delta S = \frac{\Delta Q}{T}\,. \tag{5.2}$$

Conversely, the entropy will be decreased by the same amount if the same quantity of heat were removed at that that same temperature. Entropy, like energy, is an additive quantity, but unlike energy, entropy is not a conserved quantity.

As a consistency check of the above expression, consider the example of an isolated system consisting of two bodies in contact: A large hot body at temperature T_2 is in contact with a colder large body at temperature T_1. If the hot body loses a small amount of heat ΔQ, then its entropy changes by

$$\Delta S_2 = \frac{-\Delta Q}{T_2}\,, \tag{5.3}$$

while the entropy of the colder body, which absorbs that same amount of heat, increases by

$$\Delta S_1 = \frac{\Delta Q}{T_1}\,. \tag{5.4}$$

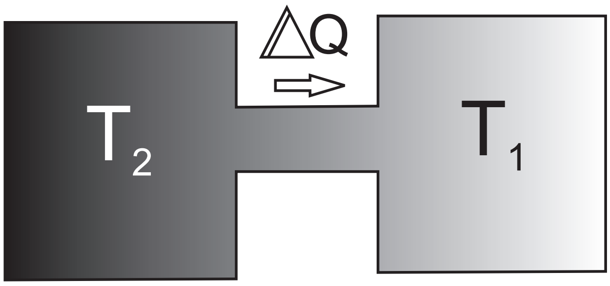


Figure 5.5: Heat flow between two large bodies

The entropy change of the whole system is therefore

$$\begin{aligned} \Delta S &= \Delta Q \left(\frac{1}{T_1} - \frac{1}{T_2} \right) , && (5.5) \\ &= \Delta Q \left(\frac{T_2 - T_1}{T_1 T_2} \right) , && (5.6) \\ &> 0 . && (5.7) \end{aligned}$$

The net entropy of the system has increased because of the heat flow from the hot body *to* the colder one. Thus using the expression (5.11) gives a result that is consistent with common experience (the direction of heat flow), and the Second Law: Notice that the First Law by itself would not have disallowed a spontaneous flow of heat from the hot body to the cold body, a phenomenon contrary to common sense and the Second Law.

It is important to bear in mind that the Second Law does not disallow a flow of heat from a cold system to a hotter system — in fact that is what an air-conditioner does! But an air-conditioner requires a power input to operate, so the heat flow is not spontaneous; in other words, the system is not isolated.

5.3.3 Maximum Efficiency of Heat Engines

Many engines and power plants burn fuel, converting part of the heat energy so generated into more useful forms such as electricity or mechanical work. In practise there are losses due to friction, leakage, and so forth. But even in the most ideal case, as we have hinted above, the Second Law places constraints on the maximum efficieny that can be attained.

Consider the example of a car engine, which extracts heat energy Q_2 from the combustion of petrol at temperature T_2, and which expels some heat, Q_1, into the exhaust at temperature T_1. By the conservation of energy, the difference $Q_2 - Q_1$ is the amount of work (mechanical energy) done in driving the car. The **efficiency**, $\mathcal{E}$, is defined as the ratio of the amount of

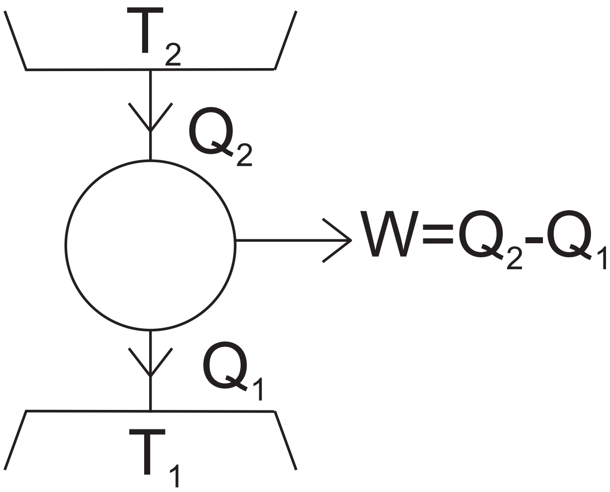


Figure 5.6: Heat Engine

useful work done to the fuel energy used,

$$\mathcal{E} = \frac{Q_2 - Q_1}{Q_2} = 1 - \frac{Q_1}{Q_2}\,. \tag{5.8}$$

However the Second Law states that the total change in entropy of the whole (isolated) system cannot decrease, that is,

$$\Delta S \quad = \quad \frac{-Q_2}{T_2} + \frac{Q_1}{T_1} \geq 0\,. \tag{5.9}$$

Combining the two equations gives

$$\mathcal{E} \leq 1 - \frac{T_1}{T_2}\,. \tag{5.10}$$

Thus the maximum efficiency of an ideal engine is controlled by the source temperature,T_2, which should be as high as possible , and the exhaust (sink) temperature, T_1, which should be as low as possible. Of course in practise T_1 can only be as low as the ambient temperature $\sim 300K$, so only the source temperature is adjustable. For example, for a combustion temperature $T_2 = 1000K$, one obtains $\mathcal{E} = 0.7$, which means that even in this ideal engine 30% of the input heat energy is wasted — that is, a proportion of the disordered heat energy will be unavailable for any useful work!

It must be emphasized that the less than 100% efficiency of the ideal engine above is a consequence of the Second Law. A real engine will be *even less efficient* because of internal friction, incomplete combustion and other factors.

From the above discussion it should be clear that when one hears the common statement "Conserve energy!", it is not that we are being told to preserve the First Law, which automatically holds, but rather to beware about wasteful use of high-grade ordered energy: Burning fossil fuels to run cyclical engines (that generate electricity for example)runs into the Second Law which demands that some energy be dumped as low-grade heat into a lower temperature reservoir.

The above limitations do not apply to direct (no conversion to work) applications of heat energy such as the use of solar energy to heat houses.

5.4 Optimizing Efficiency of Internal Combustion Engines

The original energy source in the engine is the chemical potential energy of the petrol. This is ordered energy but difficult to use directly and that is why the fuel is burnt. The initial spark provides energy (activation energy) to break the hydrocarbon bonds in petrol so that stronger bonded products can form with the oxygen in the air, with a net release of thermal energy that is used to perform work.

As we have seen above, the maximum theoretical efficiency that we can attain in converting the thermal energy into work is by burning the fuel as hot as possible and releasing the exhaust at close a temperature to the ambient as possible. Actual efficiencies are only about 30% because of various losses and imperfections.

As most of us know, the exhaust from real engines is usually hotter than the ambient temperature. The exhaust temperature can be reduced if the gases are allowed to expand as much as possible before being expelled. Thus one desires a engines with a **large compression ratio**: The ratio of the volume occupied at the end of the power stroke to that occupied at the start of the power stroke. Normal engines have compression ratios of about $10:1$, while more efficient engines have a ratio of $15:1$.

However if the fuel-air mixture is compressed too much it might overheat and self-ignite prematurely causing "knocking": Energy is wasted as the smooth operation of the cycles is disrupted. One solution is to use fuels which have a higher ignition point (high octane fuels), but most cars will work fine on regular fuel.

5.5 Air-Conditioner

While a heat-engine converts thermal energy into useful work in a system, an air-conditioner uses an input of work to remove heat energy from a system. The Second Law is of course not violated because although heat is made to flow eventually from a lower temperature to a higher one, the flow is

not spontaneous. Stated differently, when the whole isolated system is considered, the entropy actually increases as expected: The entropy of the outside where heat is dumped increases more than the reduction of entropy indoors.

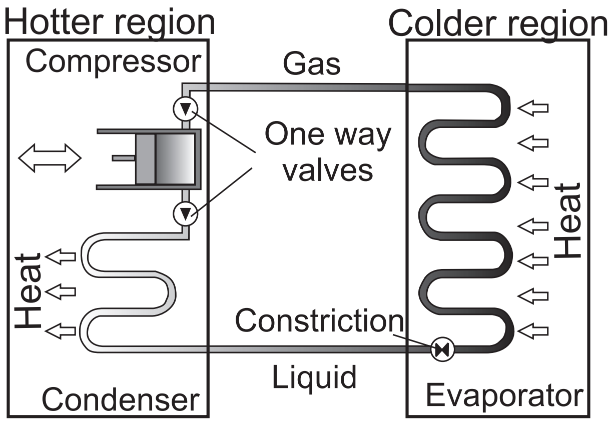


Figure 5.7: Airconditioner

Modern air-conditioners use hydrofluorocarbon (HFC) based fluids for the coolant in contrast to the ozone depleting chlorofluorocarbons (CFC's). Interestingly, CFC's were invented by scientists in the early part of the century as a substitute for toxic refrigerants such as ammonia. The advantages of CFC's were their non-toxicity, and relative inertness — they can be breathed without harm. Thus they were used as a propellant for hairsprays, deodarants and other industrial applications. Now with the dangers of CFC's known, they are being replaced by HFC's — but these too may have their problems — they may add to the greenhouse effect!

The cooling fluid has the property that its a gas at low pressure and a liquid at high pressure. This fluid flows in tubing that goes through three distinctive parts of the air-conditioner: The evaporator, the compresor and the condensor. The evaporator is located indoors and is the place where heat is extracted by the fluid from the room, while the compressor and condensor are located outdoors where the fluid gives up its heat.

When the fluid passes from the condensor to the evaporator, it goes as a liquid through a narrow portion of the pipe that causes the a drop in pressure. The liquid therefore evaporates and expands into a gas and cools as a result. This cold gas is then able to absorb heat from the room through the evaporator fins. The gas next passes to the compressor which pressurises the gas and consequently heats it up even more. The hot and dense gas next flows to the condensor which helps transfer the heat from the gas to the outside air, thus cooling the dense gas and transforming it into a liquid again, ready for the next cycle.

Did you know that modern air-conditioning was the result of trying to solve the problem of humidity on colour printing ? See Ref.[**??**]

5.6 Our Energy Sources

Given our heavy dependence on energy to derive material comfort, should we be worried about our future energy sources? Are all energy sources equally efficient? Are there any environmentally clean fuels ?

The main sources of energy currently available for human consumption are

1. Fossil fuels
2. Solar energy
3. Nuclear energy
4. Hydroelectric power

5.6.1 Fossil Fuels

Fossil fuels are carbon-rich natural deposits, formed from the remains of dead organic matter over long periods of time under high temperatures and pressures. Currently the fossil fuels coal, oil and natural gas account for about 90% of the industrialised world's energy needs.

However given the long time span of a few million years for the generation of fossil fuels, this energy source is essentially **non-renewable**. In the not too distant future – by the time fossil fuels have run out (estimates suggest oil will run out in about 100 years)– uranium could provide the requisite energy. For the even more distant future, clean energy in the form of sunlight might be the best option.

Fossil fuels are extremely polluting. Their combustion releases sulphur and nitrogen oxides which combine with the moisture in air to form "**acid rain**". Furthermore, the buring of fossil fuels increases the carbon dioxide content in the atmosphere, which traps heat and thereby causes the temperature of the Earth's surface to increase — the "**greenhouse effect**". Some believe that in the long run, this greenhouse effect will cause severe and catastrophic climatic changes.

5.6.2 Solar Energy

The energy in sunlight can be converted to other useful forms such as heat and electricity, using a several technologies. However the practicality (efficiency and cost) of this resource depends on where one is located on Earth.

The "solar constant" expresses the rate at which solar energy strikes our upper atmosphere. Since the distance of the Earth from the Sun varies

throughout the year (being shortest in December-January and longest in June-July), a mean distance of about 1.5 million kilometres is used to get the value $1400W/m^2$ for the solar constant. Obviously, not all of this energy reaches the Earth's surface. In addition to the depletion due to the atmosphere, the actual energy depends also on the latitude, longitude and local landscape of the location. All these factors affect the positioning of solar collectors.

Solar energy is not only plentiful, but it is also a clean and renewable source, unlike fossil fuels. However its direct use as an energy source is still more costly at present compared to fossil fuels. Nevertheless solar cells have been used in many practical situations, such as supplying electric power to man made satellites.

5.6.3 Nuclear Energy

Nuclear fission currently supplies about 12% of the world's energy needs, and it is likely that this source will be dominant as fossil fuels run out. The basic origin of this energy is Einstein's relation between mass and energy. In nuclear fission, an atom of Uranium splits after absorbing a slow neutron, into a number of smaller fragments. The total mass of the end products is less than the mass of the initial reactants, and the difference is precisely the produced energy.

The large value of the speed of light means that, by $E = mc^2$, even a small mass can be converted into a large amount of energy. For example, the total fission of one kilogramme of Uranium can provide the same energy as that released by burning three million kilogrammes of coal!

Nuclear fission is also the process responsible for the large energy release in an atomic bomb. The difference between an atomic bomb and a nuclear reactor is that in the latter the energy is released in a controlled manner.

A problem with energy from nuclear fission is that it is not clean, as radioactive waste products are produced which must be disposed off somewhere.

In principle one can produce energy also from **nuclear fusion**: When two atoms combine to form a bigger atom, the final product has less mass than the sum of original masses and again the difference is converted to energy. However although this is a cleaner resource than fission, and though the main ingredient, deuterium is plentiful in sea-water, fusion is far from a reality now because of various technological hurdles, and it is unclear if these will be overcome for it to be practical in the forseable future.

Amusingly though, we are the beneficiaries of nuclear fusion: The Sun produces its energy by fusing hydrogen into helium under conditions of extreme temperature and pressure. The Sun will continue to provide energy until it dies some five billion years from now.

5.6.4 Hydroelectric Power

This energy source has been used for ages but it has only been in the twientieth century that massive dams have been built in several countries to exploit their natural waterways for the production of electricity. Although it accounts for only a small portion of the world's *total* energy source, hydroelectric power is resposnible for 25% of the world's generation of *electric* power.

This source is clean and renewable, and is a conversion of potential energy first into kinetic energy and then into electricity. (The one drawback of large dams however is the ecological and sociological impact due to flooding of large areas).

5.7 Some Negative Consequences of Technological Development

5.7.1 Greenhouse Effect

The composition of the atmosphere is predominantly nitrogen (79%) and oxygen (20%). All other gases constitute less than 1% but their relative scarcity belies their importance and impact. Water vapour, carbon dioxide, carbon monoxide, nitrous oxide, and methane are the gases commonly termed the greenhouse gases because by slowing down the rate at which energy is lost by the earth, they result in a global warming effect called the greenhouse effect. Without these natural greenhouse gases, the global average temperature would be about 30C less than what it is today. However, industrial by-products like the artificial chemicals termed halocarbons such as CFCs and HFCs, sulphur hexafluoride and ozone in the lower atmosphere as well as increased release of the greenhouse gases by human activities are also contributing to what is now being called the enhanced greenhouse effect

The balance between incoming and outgoing radiation between the earth's atmosphere and outer space is called the Radiant Energy Budget and this is what drives the whole climate. According to Wien's Displacement Law which states an inverse relationship between the temperature of an ideal object called a black body to the peak wavelength it emits at, the Sun radiates in the ultraviolet and visible wavelengths due to its higher temperature. The Earth's atmosphere is transparent to these wavelengths. Therefore, incoming solar radiation (insolation) is absorbed by the Earth and re-emitted as lower wavelength Infrared Radiation (IR) because of the Earth's lower temperature. The greenhouse gases, primarily water vapour and carbon dioxide absorb certain amounts of this outgoing terrestrial radiation thereby holding it in the atmosphere for a longer period and thus allowing for a warmer earth and a more even distribution of temperature between lower latitudes which experiences a daily net positive energy budget and the higher latitudes

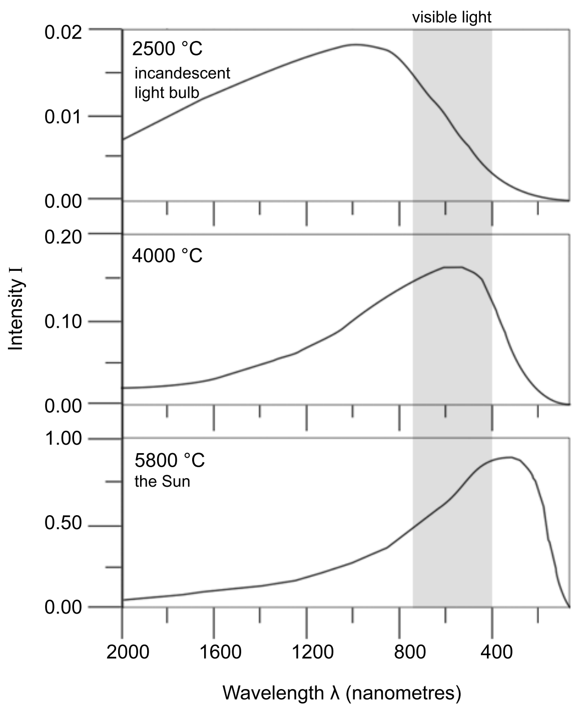


Figure 5.8: Variation of intensity and peak of radiation of a black body with temperature

which experience a negative one. The result is a higher global temperature (15C instead of -18C) as well as a lower temperature differential between the poles and the equator. Eventually however, terrestrial radiation is dissipated to space so that over the course of a year incoming radiation and outgoing radiation balance out.

As a result of human activities, this energy budget is changing upwards as the amount of all the greenhouse gases in the atmosphere, primarily carbon dioxide, increases. Outgoing radiation is being held for longer, raising temperatures globally and this is thought generally to have dramatic consequences for the earth's climate.

Though the greatest contributor to the greenhouse effect is water vapour, it can only indirectly be affected by human activities. Therefore, the phenomenon of Global Warming is mostly (60%) a result of the increased level of carbon dioxide in the atmosphere. The United Nations estimates that the atmospheric carbon dioxide level has increased by over 30% since 1800s with current annual emission level at 23 billion metric tons, which is already almost 1% of the level in the atmosphere. Global temperatures have risen by 0.6C since the last century and the general scientific consensus is that

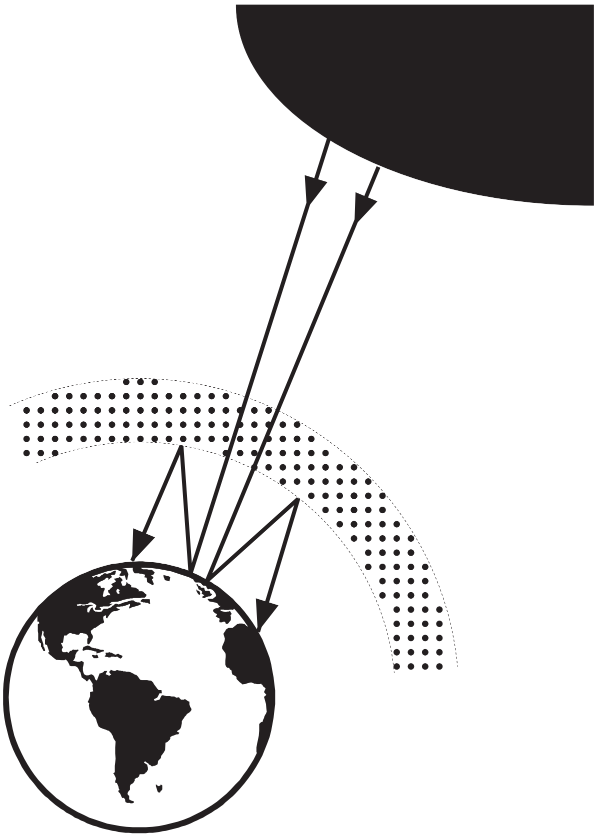


Figure 5.9: Trapping of energy by the atmosphere

there is a direct correlation between these two trends.

Normally, there has always been an exchange of carbon between the atmosphere, the oceans and land vegetation, called the natural carbon cycle, with levels of carbon dioxide having varied by less than 10% over 10 000years. The oceans are the biggest reservoirs of carbon and holds 39 000 billion metric tons while the atmosphere holds 750 billion tons. Total carbon in fossil fuel reserves contain another 5 000 billion tons and it is this previously contained store that, as a result of industrialisation, is now being released into the atmosphere as well as the 610 billion tons held in forests through deforestation. As a result of the already present rise in greenhouse gases, the global energy budget has shifted by 2.5 Watts per square metre which is an increase of 1% of net incoming solar radiation. This increase in energy is equivalent to that obtained by burning 1.8 billion tonnes of oil per minute, which is 100 times the world's current commercial energy consumption.

Therefore, it is not surprising that climate change should result. However, the alleged existent and potential consequences of climate change are

controversial as the climate is a highly complicated system with many feedback mechanisms and the timescale we are looking at is so short compared to the timescale of such global processes which normally occur over millennia. However, what is certain is that carbon dioxide levels have risen as well as global temperatures and there is also a sea level rise of 10-20cm which is consistent with the observed warming.

Some other evidence that Global Warming is indeed happening:

- 10% reduction in snow cover over the northern hemisphere since the 1960s and a thinning of the Artic ice cap

- More precipitation of $0.5 - 1\%$ in high latitudes of the Northern hemisphere coupled with reduced precipitation in the sub-tropical regions. Vulnerable areas such as sub-Saharan Africa and Asia have experienced increased frequency and intensity of droughts.

As Global Warming continues, there are likely to be consequences associated with this change in climate. Some of these predicted consequences are:

- Shifts in rainfall patterns

- Higher evapo-transpiration (loss of water from land and vegetation) rates leading to a drying out of soil during the summer planting season

- Extreme weather causing physical damage

- A mean sea level rise of 9-88cm by 2100 coupled with the melting of glaciers causing flooding of coastal and other low-lying areas

- Biosphere changes resulting in new geographic distribution of disease vectors as well as food growing areas.

- Reduction in sources of fresh water

- Continued rise in temperatures of 1.4-5.8C by 2100

The threat to fresh water and food supplies and the massive migration of people that drought and flooding would entail are the most serious consequences of global warming.

If carbon dioxide emissions remain unchecked, it is predicted to rise from today's level of 367 parts per million (ppm) to 490-1260ppm by 2100, which would be a $75 - 350\%$ increase since 1750. However, even if emission levels are checked now, the effects of Global Warming would continue to be felt as the climate is only now responding to past emissions because of the graduating influence of the oceans.

The primary cause of this anthropogenic carbon dioxide is the rapid rate of industrialisation that has been taking place since the 1800s. Power generation, automobiles and industries are the specific culprits. Deforestation to accommodate expanding populations, supply land for agriculture and produce timber also contribute by releasing carbon stored in trees and by preventing atmospheric carbon from otherwise being removed by trees.

Agriculture is the source of nitrous oxide, which is produced by fermentation in rice fields, and methane, which is from cattle. Halocarbons like CFCs and HFCs are used in many industrial products and processes like packaging and refrigeration.

The biggest carbon emitter is the US at 24% (1998 figures) of the world total followed by China at 13%. However, these figures can be deceiving. If we look at per capita carbon emission, then the US stands at 19.7 metric tons and China at 2.7. This comparison raises the most controversial issue in curbing carbon emissions: that while the majority of present damage has been caused by the developed countries on their early industrialisation drives, their carbon emissions though the highest in terms of per capita are now actually stabilising. Whereas, developing countries in their turn industrialising and raising the living standards of their citizens will be impeded by these measures. Moreover, the greatest disasters of Global Warming are going to affect the poorest and most vulnerable while the richest countries have the economic and technological resources to mitigate any effects they might experience.

Though political solutions such as the 1992 Rio Convention and the more comprehensive 1997 Kyoto Protocol are necessary to persuade countries to attempt carbon emission level reductions, ultimately and ironically it lies with technological solutions to actually reduce the carbon levels. Some solutions are to:

- Increase efficiency of power generation, automobiles and industrial processes. Indeed, this has already happened in response to the 1973 oil crisis so that now less energy is needed to produce each GDP dollar in the US and Europe.

- Seek alternative, low carbon-intensive fuel sources such as hydrogen fuel cells and biofuels.

- Minimise pollution from fossil fuel extraction.

Moreover, technology would be needed to combat the expected consequences of Global Warming such as in constructing better flood barriers, improving the productivity of agriculture to compensate for a smaller farmland supply and controlling disease vectors.

References:

- United Nations Framework Convention on Climate Change and the Kyoto Protocol.

- Oklahoma Climatological Survey

- Kasting, James F. The Carbon Cycle, Climate, and the Long-Term Effects of Fossil Fuel Burning. US Global Change Resource Information Office (GCRIO) for Global Change Resource Programme (USGCRP), Columbia University.

5.7.2 Ozone Depletion

Ozone is a special type of oxygen molecule which, instead of containing two atoms of the element, possesses three. Ozone is created by the chemical reaction between incoming solar radiation and oxygen in the atmosphere and is primarily found in the stratosphere, a layer of the Earth's atmosphere, at a height of 10-50km above the surface of the Earth. This layer of atmosphere experiences a temperature inversion where temperature increases with altitude as opposed to the troposphere, the lowermost layer. This is a result of the presence of ozone which absorbs radiation in the ultraviolet wavelengths thus heating up the stratosphere and also protecting life on Earth's surface to which this radiation is deleterious.

Chlorofluorocarbons, a group of synthetic chemicals, were invented in 1928 for use as refrigerants, propellants and packaging materials and were initially highly favoured as they were non-toxic, non-corrosive and non-flammable. However, these long-lived chemicals get broken down into chlorine in a reaction with sunlight in the stratosphere where they act as a catalyst and break down ozone to devastating effect. A single chlorine atom has the potential to destroy 10 000 ozone atoms.

Due to gravitational forces, the Earth is not a perfect sphere but rather bulges out at the equators and is flatter at the poles. Therefore, the atmosphere over the poles is thinner than elsewhere with corresponding lower concentrations of ozone and consequently, particularly susceptible to ozone depletion. The ozone hole over the Antarctic, discovered in 1985 by British scientists, is an area of low ozone density. There is a similar, though smaller one, over the Artic as well.

The resulting increase in ultraviolet radiation reaching the Earth's surface has a detrimental effect on living cells. In humans, it is linked to increased incidence of skin cancer. The shorter wavelength, UV-B is especially harmful in this respect. The American Environmental Protection Agency (EPA) predicts that if Ozone depletion continues unchecked there will be an additional 155 million incidents of skin cancer over the next century. There is also the effect on ecosystems to consider since increased amounts of UV radiation decreases the yield of plants and are harmful to animals as well. Especially, UV reduces the extent of delicate phytoplankton in the oceans on which the whole ocean ecosystem depends.

Since the discovery of the ozone hole and the linkage with skin cancer, various concerted international efforts have resulted in the Montreal Protocol (1987) phasing out the use of CFCs through substitution with alternatives such as halocarbons (HFCs) and the eventual ban of all use of CFCs as well as other ozone depleting agents such as methyl bromide and halons. This has been largely an environmental success story. However, the past production and use of CFCs continue to extol a penalty since due to the temperature inversion pollutants are removed from the stratosphere at a very slow rate.

Reference:

- Brennan R. Levitating Trains and Kamikaze Genes. Wiley, 1990.
- The United Nations Environmental Programme, The Ozone Secretariat
- US Global Change Research Information Office (USGCRIO)

5.7.3 Acid Rain

Acid rain or more accurately, acid deposition/precipitation refers to normal precipitation that has been rendered even more acidic as a result of pollutants. It also refers to dry deposition of the pollutants in the air. Normal precipitation (rain, snow, sleet, fog) is already slightly acidic with a pH of 5.6. The presence of pollutants such as sulphur dioxide and nitrogen oxides

from the combustion of fossil fuels in the atmosphere increases this acidity even further.

The primary causative factors are power generation, especially coal burning power stations and vehicles. The effect of increased acidic precipitation causes damage to vegetation such as defoliation and even their destruction. The runoff acidifies freshwater streams and lakes making them inhabitable for fish and other water life.

Acid precipitation is primarily a regional and local phenomenon endemic to areas of high concentrations of industries. However, atmospheric circulation patterns do enlarge the susceptible areas. Therefore, for example, pollutants from America's industries in the North affect Canadian forests and fisheries.

The same measures needed to reduce Greenhouse Gas emissions will also have the salutary effect of reducing the incidence of Acid rain. Burning clean coal which is low in sulphur or treated coal that has had the sulphur and other contaminants removed would also greatly reduce the amount of pollutants released into the air. Cleaner and more efficient vehicles would also have the same effect.

5.8 Reduce, Reuse, Recycle

In addition to energy conservation, it is clearly desirable to prevent waste of useful materials. Reducing waste and reusing what can be reused (perhaps for an alternative purpose) are common and economically sensible activities in most developing countries. In many industrialised countries however these practices had almost disappeared among the affluent population and have been re-emphasized now because of the concern for the environment. If it is not possible to reduce or reuse then one attempts to recycle and many companies are now trying to redesign their products to make recyclng easier. However as many of the 3R activities are often seen as time consuming and not economically attractive, their acceptance requires a re-education of the inhabitants of this planet.

5.8.1 Waste Disposal

Waste disposal is becoming a huge environmental problem as more of the world industrialises and people become more affluent with the consequent rise in waste generation. However, land is limited and even then, not all waste can be landfilled or incinerated.

The problem is particularly acute as regards certain types of poisonous or toxic waste. Such waste is the product of certain industries such as the petroleum industry or paper mills. This waste is difficult to handle and cannot be simply landfilled as it will seriously pollute the soil and underground water with consequent harm to public health and incineration only

produces other harmful waste such as poisonous gases and ashes. Most solutions consist merely of storing such waste. However, technology here has an important part to play. New methods of incineration are being developed such as vacuum extraction which is effective for volatile substances. Biotechnology is being used to develop special microbes that eat organic waste and degrade it into less harmful by-products though this has its limitations.

Nuclear waste is the most difficult and costly of all waste to handle. Because it is radioactive, it is very dangerous to all living things and requires extremely cautious handling and disposal. They are also very long lived, being dangerous for millions of years. They are the result of nuclear power and weapons production. Nuclear waste is so-called because it has become too radioactive for safe handling and therefore, not viable for further use. The only current solution has been to seal it off in bulky lead and concrete containers and bury them deeply in remote locations where it is hoped that the intervening layers of rock and soil will isolate the radiation. However, this is only a temporary solution and locations for burial are in short supply whereas the worlds' current nuclear power stations and weapons production facilities continue to generate nuclear waste.

5.8.2 Waste Management and Recycling in Singapore

Singapore is divided into four areas for the purpose of waste collection with each being serviced by its own private waste collection company. For example, the Woodlands-Yishun area is serviced by Semac, a subsidiary of Sembcorp and is the largest such service provider. Due to the nature of collecting domestic waste existent in Singapore, little or no sorting for the purpose of recycling is done at the source, although this is changing quickly. Collected waste is generally sent to one of the four incineration plants in Singapore of which the Tuas South Incineration Plant (TSIP) is the latest.

Incineration is the solution sought in Singapore as incinerating waste reduces its volume by up to 90%, thereby reducing the amount of waste that needs to be landfilled. The Tuas Incineration plant is the fourth incineration plant in Singapore built at a cost of 900 million Singapore dollars and completed in 2000. It handles both domestic and industrial waste. No sorting of the waste is attempted before the incineration process but after incineration the ferrous material is removed by magnets and the remaining ash and slag is sent by barque to an off-shore landfill at Pulau Semakau, though some of the ash is used in road construction. Electricity is generated as a by-product of burning of which the plant consumes 20% and sells the rest to Singapore Power. The ferrous material is sold to local smelters. Through these means, the plant is able to cover its running costs and even enough make a small profit.

The TSIP is the most high technology plant to be constructed in Singapore with numerous features designed to increase efficiency, reduce pollution

and lower operation costs. Its maximum capacity is 3000 tonnes and it is in fact the largest such plant in the world. It currently incinerates 2500 tonnes of waste daily and is therefore, approaching maximum capacity already. The non-sorting of waste before incineration poses a problem as it results in uneven combustion, which is mitigated by thorough mixing, and the incineration of toxic materials such as batteries. However, the plant possesses 150m high chimney flues which are equipped with technologies such as electrostatic precipitators to remove harmful emissions and minimise pollution of outlying areas thus confirming to the strictest environmental standards. To conserve water, TSIP uses industrial water that it pre-treats itself for use in the boilers that generate electricity and for emergency use in case of fires.

The ash from the 73% of waste that is incinerated and the non-incinerated waste are landfilled in Pulau Semakau, the last such site in Singapore and after the closure of the landfills based on the mainland, the only one left. It was commissioned in April 1999 and constructed at a cost of 610 million singapore dollars by building a 7km long rock bund to enclose an area of the sea around two islands, Pulau Semakau and Pulau Sakeng. This area is further divided into cells sealed off by impermeable membranes and marine clay to prevent any leachate from polluting the sea. The total area is 350 hectares and has a capacity of 63 million cubic metres. The waste is offloaded and transported by trucks to the tipping sites and as each cell is filled, it will then be covered with earth allowing for plants to take hold later. Mangroves have been planted around the perimeter as pollution detectors. After all the cells have been filled, the island may be used for future industrial or recreational needs.

Singapore faces a waste management crisis with TSIP already operating at close to maximum capacity and the projected life of the Pulau Semakau landfill at only 20 years with no other alternative available when it becomes full. Therefore, recycling programmes have been initiated by the Environment Ministry. One such programme to promote recycling among households entails giving them a special plastic carrier in which to place all possible recyclable items. These bags are collected individually from each household by private recycling companies who have been sub-contracted this task by the respective waste collection company for that area. The collected recyclables are then transported to a sorting centre, the SembVisy Materials Recovery Facility, a collaboration between Sembcorp and an Australian company Visy Recycling. Private recycling companies abound which recycle industrial waste. Some other companies collect recyclable items from public places. An example would be Workmax Engineering which collects drink cans from canteens and food courts, sorts and compacts them for recycling. Some items such as ferrous metal are recycled in Singapore while others such as aluminium are exported to other countries like China for recycling.

The 2012 targets are to increase the recycling rate in Singapore from the

current 44% to 60% among households and thereby extend the shelf life of Pulau Semakau to 50 years as well as extend the rate of incineration plant building from the current one every 5-7 years to one every 10-15 years.

5.9 Evolution of the Aluminum Can

Designing a product is more than a feat of engineering alone. Economic, aesthetic and environmental factors play as important a role as physical possibilities. This fact is amply demonstrated by the evolution of the aluminum can, which though seemingly simple and commonplace, is a remarkable achievement of product design.

The early cans used to store food were made of iron and were extremely heavy and difficult to open, literally requiring a hammer and chisel. With the development of steel, cans became lighter but remained as difficult to open as before.

Canning carbonated beverages posed a special case as since they possessed an internal pressure, the cans need not possess great strength and so the can walls can be made thinner. Aluminum has less strength than steel but is much lighter and more ductile. Therefore, it allows for seamless construction that makes for relatively cheap mass production. The entire body of the can except for the top is made from a single sheet of aluminum. However, since its inception in the 1960s, the ubiquitous aluminum can has undergone extensive evolution for manufacturing, design and environmental reasons.

The most important improvement from the manufacturing point of view was cost reduction and this involves reducing the amount of aluminum needed to make one can, in a process called 'lightweighting'. The top of the can contains the most material so to reduce its size, the can has become narrower and taller as much as possible without losing too much of a can's proportions. The top was made even smaller through a slight tapering. As a result of lightweighting, today's aluminum can contains 40% less material. However, a can needs to possess a certain minimum strength to allow for handling, stacking during transportation and to resist the internal pressure of the liquid so there is a minimum thickness. The top has to be stronger than the walls to allow for the opening tab. The bottom of the can is dome shaped to take advantage of the strength of the arch to prevent the weight of the liquid from making it bulge out but without the need to reinforce it.

The greatest design evolution has concerned the self-opening tab. Early cans as mentioned before needed a can opener or a special church key provided with the can to open it. This was complicated and inconvenient in cases when can openers were not to be found. A certain Ermal Fraze then invented the pop top, the first self-opening can. Even after ironing out certain engineering problems involving making the pop top strong enough so it

does not spontaneously open while making it easy for an user through careful scoring of the top and ensuring that it does not leak, certain environmental problems cropped up. The pop top was removable, leading to a major littering problem. Being a small, sharp piece of metal also meant a major safety problem when people who dropped it into their drinks swallowed it accidentally, people walking with bare feet stepped on it and children picked it up and swallowed it. Therefore, the non-removable tab as we know it was invented by Daniel F. Cudzik in 1976 which solved all these problems and helped to recover an amazing four million tons of aluminum that would otherwise have been lost when aluminum can recycling took off in a big way.

Aluminum is very expensive to produce, even more expensive than steel, as electrolytic reduction is needed to extract aluminum metal, thus requiring enormous quantities of electricity to do so (at least 12 kilowatt hour per kilo of aluminum). However, it is very cost effective to recycle aluminum cans since it is entirely made of just one metal and even the paint used in the packaging aids the process of smelting through providing volatile gases that can be used as fuel. Recycling aluminum requires only 5% of the energy needed to make new aluminum, a saving of 14 kilowatt hours of electricity per kilo. Moreover, there is no downgrading involved in recycling aluminum so it can be reused in any way. Therefore, recycled aluminum is a reliable and cheap source of the metal and a recycling success story (in 2000, about 862 million kilo or 60% of aluminum cans were recycled in the US with recycled aluminum constituting 35% of the industry's supply). Recycling aluminum saves not just resources and energy but pre-empts the environmental pollution inherent in landfilling and new production. The whole process has become so streamlined that it takes only six weeks for a dumped aluminum drink can in the United States to become a brand new can of drink on a supermarket shelf.

The evolution of the aluminum can has not reached a stasis yet as engineers continue to improve upon it. In the mid-1990s, a new innovation called 'fluting' appeared which made the cans look like classical Roman pillars. This design made the can lighter by 10% while increasing its strength by 20%, which should make it a successful design. However, it proved unpopular with the fickle customer who preferred the traditional contours. In the mid-1990s, Coca-cola pioneered the curvy can which evokes the classical glass coke bottle for aesthetic reasons only since this design actually uses more material. However, it is easier to hold and contains the pressure better. It remains to be seen if this will take off (it is still being tested by Coca-cola).

Fact Sheet:

- Through 'lightweighting' which includes narrower walls, tapering and other refinements, the amount of material in an average can has been

reduced by 40% since the 1960s.

- It takes about 14 kilowatts hour per kilo of electricity to produce new aluminum but recycled aluminum requires only 5% of this.

- Aluminum is the only profitable recyclable and recycling aluminum does not result in a downgrading of quality.

- Recycling aluminum cans saves energy and other resources; reduces pollution from not having to increase aluminum extraction and not having to landfill used cans and creates new business in recycling.

- 60% of all aluminum cans in the US and a higher percentage in certain European countries (92% in Sweden) is recycled.

References:

Aluminum Association. www.aluminum.org ¡11 June 2003¿

Alcoa (aluminum producer) http://www.alcoa.com/global/en/home.asp ¡11 June 2003¿

Henry Petroski. Invention by design : how engineers get from thought to thing USA: Harvard University Press, 1996.

International Aluminum Institute. www.world-aluminum.org ¡18 June 2003¿

5.9.1 Packaging for Fast Food

The success of a good design depends on more than just its form and function. Technical details are just one part of the story with the rest made of up social and other factors which evolve with the times.

In the early 1970s, MacDonald's wrapped its burgers with paper. However, the newly introduced Big Mac required more careful packaging and this was accomplished in a complicated manner through stiff paper collars and a double wrap of foil and paper. It was complicated and time-consuming to pack and unpack and did not suit the spirit of a fast food company. So when the polystyrene clamshell box was introduced, everyone felt that this was a vast improvement. A single packaging material that was able to accomplish

various functions perfectly: hold the burger's shape, maintain the heat, absorb the grease, simple to pack and unpack and it was good advertising for the new product with a suitably colourful box.

However, it had such a short lifespan that it seemed a wasteful use of resources. And it was bulky and non-biodegradable as well as non-recyclable and therefore, came to constitute "the most visible symbol of profligate disregard for the environment" (223). MacDonald's was responsive to consumer opinion. When the case was made that the CFC in the packaging was hurting the Ozone layer, it quickly switched to non-CFC ones. However, this was not enough to satisfy the environmentalists and eventually it was decided to go back to paper even though the original reasons for the change remain. Paper is not ideal as well since it is also a pollutant and is non-degradable too; however, its effect on the environment is not as extensive as that of polystyrene boxes.

This very public example goes to show that a good design is much more than simply about technology. Various factors such as in this case, consumer's concern for the environment, can come to play a crucial role in its success. What has become of consideration now is whether designers anticipate the impact the product is likely to have and include this in the conceptualisation process.

Reference

Petroski, Henry. The evolution of useful things

5.10 Summary

- Heat is a form of energy, it is the kinetic energy of atoms and molecules which are in constant motion.

- Laws of Thermodynamics
  1. First Law (Law of Energy Conservation): The change in internal energy of an object is the difference between the heat input and the work done by the object
  2. Second Law Statements:
  (a) Kelvin's: A cyclical engine that transforms all heat completely into work is not possible.
  (b) Clausius': An engine cannot solely transfer heat from a colder reservoir to a hotter one (Heat does not flow spontaneously from a cold object to a hotter one).

- Entropy
  1. Entropy is a measure of the disorder in a system. By the Second Law, the entropy of a closed system never decreases.

2. The change in entropy caused by adding a very small quantity of heat slowly (reversibly) to a system at fixed absolute temperature T is given by

$$\Delta S = \frac{\Delta Q}{T} \, . \tag{5.11}$$

- The efficiency of an engine is the ratio of the amount of useful work done to the input heat energy.

- Energy Sources
1. Current energy sources constitute those from fossil fuels, nuclear power, hydroelectric power, solar energy and biofuels. 2. Since our major energy sources are non-renewable and finite in quantity, we need to economise on high-grade (ordered) forms of energy and it is also advisable to reduce wastage of resources since energy is involved in their production.
3. Many fuel sources are polluting and result in other negative effects such as the greenhouse effect and the ozone problem.

5.11 Exercises

1. In Boltzmann's time, many people did not believe in atoms. Apparently Boltzmann, who gave the molecular interpretation of entropy, became depressed by the many attacks on his work and eventually took his own life in 1906; [5]. Are you convinced that atoms exist? Check out the pictures in Ref.[4].

2. The Law of Conservation of Energy states that energy cannot be destroyed but only converted from one form to another. Then what does one mean by statements such as "Do not waste energy" or "Conserve energy" ?

3. (a) What is the entropy change when 1 kg of water is converted to steam at 100 degrees Celcius?
(b) What is the entropy change when 1 kg water is converted to ice at 0 degrees Celcius?
(c) Do the signs of the entropy change above agree with increase/decrease of order in the system ?

4. A refrigerator works by extracting heat Q_1 from the interior at temperature T_1, and expelling heat Q_2 to the outside at a higher temperature T_2.
(a) Show that conservation of energy, implies that the electrical power

source must supply work equal to the difference $Q_2 - Q_1$.
(b) The **efficiency**, $\mathcal{E}_r$, of the refrigerator is defined as the ratio of the amount of heat extracted to the work supplied. Show that the Second Law constrains the efficiency to be

$$\mathcal{E}_r \leq \frac{T_1}{T_2 - T_1} . \tag{5.12}$$

5. (a) Why can you not cool the kitchen by leaving the refrigerator door open ?
(b) What is the difference between the action in (a) and the use of an air-conditioner?
(c) Compute the maximum efficiency of the air-conditioner you use at home.
(d) Compare your result in (c) with the stated efficiency of the air-conditioner.
(e) How much heat is extracted from the room by your air-conditioner for every Joule of electrical energy used ?
(f) Where does the heat extracted in part (e) go ?

6. When water in a glass is frozen in a refrigerator to form ice, the amount of disorder of the system clearly decreases when one considers the motion of the molecules. Thus the entropy of the glass of water decreases. Does this violate the Second Law ? Explain.

7. Suppose you turn the key, the engine "turns" but the car does not start. What could have gone wrong ?
(Hint: See http://www.howstuffworks.com/engine3.htm)

8. A diesel engine does not use spark plugs. How does it ignite the fuel? What are the advantages and disadvantages of a diesel engine compared to a petrol engine ?

9. How does the catalytic converter of a car help to reduce pollution ? (Hint : See [6]).

10. (a) Fuel cells convert chemical energy directly to electrical energy and are used in spaceships. Why are they more efficient than combustion engines ?
(b) Can fuel cells be used in cars ?
(c) What other fuel can be used to power cars ?

11. Make a list of energy sources humans have developed and discuss which are the most widely used and which are the most environmentally friendly.

12. (a) Give four examples of renewable energy sources and describe their origins.
    (b) Give four examples of nonrenewable energy sources and describe their origins.
    (c) Which of the renewable energy sources above are practical for our nation ?

13. It is sometimes stated that "wind energy is another form of solar energy".
    (a)Explain the above statement.
    (b) Give practical examples of the harnessing of wind energy.
    (c) In what sense are hydroelectric power and wind power forms of solar energy ?

14. Read the relevant sections of this chapter and/or other sources and then answer the following questions.
    (a) How do the naturally occuring carbon dioxide, methane and water vapour in the atmosphere warm our planet ?
    (b) How have human activities increased the percentage of carbon-dioxide in the atmosphere?
    (c) Which are some of the other greenhouse gases produced by human activity ?
    (d) What might be the long term consequences of slight global warming caused by the greenhouse effect ? Has there been any evidence of this?
    (e) What is ozone, where does it occur naturally, and what important role does it play ?
    (f) Why is the presence of ozone bad in the lower atmosphere bad ? How does it arise there?
    (g) What man-made products are most likely to be responsible for the ozone hole ? Where are those man-made products used ?
    (h) What is acid rain, how is it caused, and what are its consequences ?
    (i) What can you do personally to alleviate the problems mentioned above?
    (j) What can we do as a nation to alleviate the above problems ?

15. Read the sections on aluminium can recycling and fast-food packaging for classroom discussion.

16. (a) What are the benefits and costs of the 3 R's : Reduce, Reuse and Recycle?
    (b) Read the section on waste management in singapore for classroom discussion.

17. (a) It is sometimes stated that technology itself can help us solve the problems created by technology. Give some examples of this with respect to the problems of the last three question.
(b) It is also said that every technological solution creates its own new problems. Do you agree ? Illustrate your position with some examples.

5.12 References

1. Brennan, R., Levitating Trains and Kamikaze Genes, and
Ball,P. , Designing The Molecular World, Chemistry at the Frontier.

2. Petroski, H., The Evolution of Useful Things, pg 220-225.

3. Petroski, H, invention by Design, Chapter 5 Aluminium Cans and Failure

4. Atoms are real. See for example
http://www.fourmilab.ch/autofile/www/section2_84_14.html

5. Boltzmann at
http://www-groups.dcs.st-andrews.ac.uk/∼history/Mathematicians/Boltzmann.html

6. Catalytic Converter at The How things Work Website
http://www.lenoxsoftworks.com/cgi-bin/htmlos.exe/00235.5.1611683706100009720?

7. Fuel Cells at http://www.howstuffworks.com/fuel-cell1.htm

Chapter 6

Electromagnetism: Motors and Generators

I have not failed 700 times. I have not failed once. I have succeeded in proving that those 700 ways will not work. When I have eliminated the ways that will not work, I will find the way that will work.
–Thomas Edison

It was known by the 18th century that by rubbing together different materials one could generate two types of charges, positive and negative, with like charges having the property of repelling each other while unlike charges attracted each other. It was also learnt how to store these charges in devices called Leyden jars and that these stored charges could be used to create sparks.

Although many people noted the similarity between the man-made sparks and lightning, but surprisingly, superstitions concerning the origin and cause of lightning persisted up till the beginning of the 18th century because of the lack of a systematic scientific investigation. It was Benjamin Franklin who first investigated the phenomena of lightning using the **scientific method** of comparing hypothesis and theory with experiments. Franklin first hypothesised that clouds were electrically charged, and if this were correct, then lightning would be easily understood as an electrical discharge similar to those seen in laboratories.

To test his hypothesis, Franklin flew a kite during a thunderstorm in Pennsylvania in 1752. He apparently took the precaution of standing under a shed, and holding the kite string with a dry (insulating) silk cord. (Some experimenters at that time had been struck by lightning and killed!). When the string became wet, the electric charge accumulated by the kite flowed down it to a metal key attached to the end. Franklin observed a spark jump between the key and his hand. Thus Frankin had demonstrated the electrical nature of thunderstorms!

Franklin went further: By touching the key to a Leyden jar, he charged it and used it to deduce that the lower part of thunder-clouds was usually negative! Franklin did many other experiments, and also invented useful devices such as the lightning rod that saved many buildings.

In the following sections we will first summarise the modern understanding of electricity in general and then discuss some applications.

6.1 Electricity

6.1.1 Charges

We know, after years of experimental and theoretical study, that matter at the microscopic level is made up of building blocks called atoms. A crude model of the atom is that of a number of electrons orbiting a core called the nucleus. The electrons are held in their orbit **not** by the gravitational force which is too weak at that scale, but by electrostatic force of attraction between unlike charges: The electrons are negatively charged while the nucleus is positively charged.

A neutral atom has equal amounts of positive and negative charges, but because it is relatively easy to strip the electrons from the outer orbits, one very often ends up with charged atoms. Indeed this is what happens when one rubs different materials together to form charged objects.

Coulomb studied the forces between electric charges and came up with a law which is quite similar to Newton's law for gravitation. Newton's law states that the magnitude of the gravitational force between two masses M_1 and M_2 separated by a distance R is given by

$$F_g = \frac{GM_1M_2}{R^2} \tag{6.1}$$

where $G = 6.7 \times 10^{-11} m^3/(kgs^2)$ is the gravitational constant, and the *direction* of the force is along the line joining the two charges.

Similarly, the magnitude of the force between two point charges Q_1 and Q_2 separated by distance R in vacuum is given by

$$F = \frac{k_c Q_1 Q_2}{R^2} \tag{6.2}$$

where $k_c = 9 \times 10^9\ Nm^2/C^2$ is a universal constant, and the *direction* of the force is along the line joining the two charges. *Notice an important difference though between Coulomb's law and Newton's law of gravitation: While gravitational force is always attractive, as there is only one "type" of mass, electrostatic force can be both attractive or repulsive depending on the signs of the charges.*

The unit of charge is the **Coulomb** (C). The charge on the electron is $1.6 \times 10^{-19} C$. It has been determined that electric charge comes in **discrete**

units, with the electron charge as the basic block, and that **charge is conserved**: That is the total charge in a closed system is fixed.

6.1.2 Fields

Both Coulomb's law and Newton's law of gravitation describe *action at a distance*. The forces between objects do not seem to require an intervening body and seem to act instantaneously. However although this was considered bizarre already by Newton, it was not until much latter that a satisfactory formulation appeared.

Until the mid-nineteenth century, all material entities were thought to be similar to a solid body: That is, occupying a definite position in space, and possessing energy and momentum.

A major conceptual revolution was brought about by physicists such as Michael Faraday and James Clerk Maxwell in the nineteenth century. In their study of electricity and magnetism, they discovered that the effects of for example an electric charge could be best understood by postulating an entity called the **field** that carries the effects of the charge. So, for example, if we introduce an electric charge into a space, the electric charge generates an **electric field** which then acts on say other charges. If the charge introduced is a positive charge, the field will act on other neighboring charges and create an attractive force on negative charges and a repulsive force on positive charges and so on. **Thus there is no action at a distance in the field picture: All forces are local**.

In contrast with a particle which occupies a definite position in space, a field is spread **all over space**. While a particle is described by specifying its velocity and position in space, the electric field, for example, is described by assigning numerical value for the field **at every point in space**. In other words, a field is a function of space (and of time), and has much more information than say a particle. Just like a particle, a field has energy, momentum, and angular momentum.

The field hypothesis is not just an empty abstraction or only an aid in visualisation. It has dramatic consequences such as the prediction of propagating waves that arise from disturbances in the field, as we shall see latter.

Mathematically, the electric field produced by a point charge is defined as the force it produces on a unit positive test charge. Therefore, an electric field $\mathbf{E}$ exerts a force $\mathbf{F}$ on a charge Q given by

$$\mathbf{F} = Q\mathbf{E} \tag{6.3}$$

Thus the SI unit for electric field is Newton/Coulomb or **N/C**.

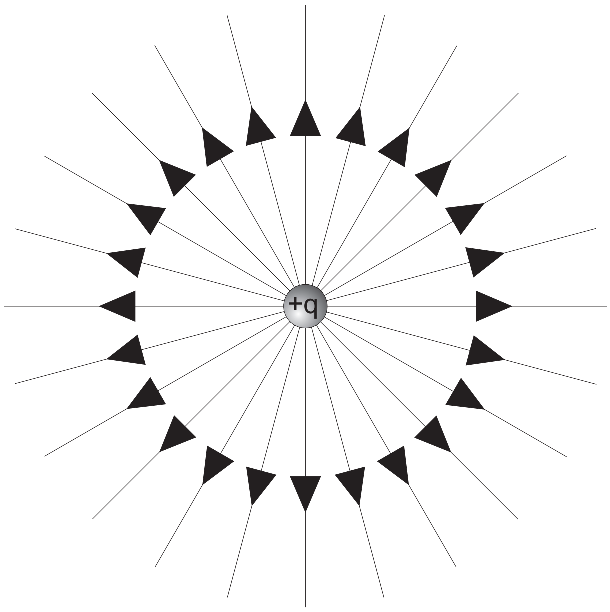


Figure 6.1: Electric field due to a positive point charge

6.1.3 Potential

In an earlier chapter, we encountered the concept of *potential energy*, an example being that gravitational potential energy — it requires work, done against the force of gravity, to lift an object a certain height, and the object is then said to have gained potential energy. Instead of talking about the energy of the object in the gravitational field, one can introduce the concept of a *potential field*: One defines the **gravitational potential** at each point, relative to some reference zero, as the *potential energy gained by a unit test mass* moved to that point.

Thus in a gravitational field a free particle moves from a point of higher potential to one of lower potential, and the kinetic energy gained is (the difference in potential)×(mass of particle).

Similarly, instead of talking of electric forces and fields which are vector quantities, one can introduce the concept of **electric potential difference** between two points, *defined as the work required to move a unit (one Coulomb) positive test charge between those two points.*

Thus the unit of electric potential is Joule/Coulomb and this is simply called the **Volt**, (V), in honour of Alessandro Volta, the inventor of the first **battery**.

Note that in an electric potential field, a free postively charged particle will move from a point of higher potential to one of lower potential, while because of the sign difference, a negatively charged particle will move in the

opposite direction.

6.1.4 Batteries

A battery is a device that creates potential differences through chemical means. Chemical potential energy in the battery is used to push electrons from the positive terminal to the negative terminal and this continues until the electrostatic potential difference becomes large enough for equilibrium to be reached. At this point, small dry cells have a potential difference of 1.5 V between their two ends.

Some batteries are rechargeable, meaning that once the chemical energy is used up and no longer able to maintain the potential difference, passing a current in the opposite direction allows the battery to rebuild its chemical energy and thus its potential. However most common batteries are **not** rechargeable and can explode if an attempt is made to recharge them.

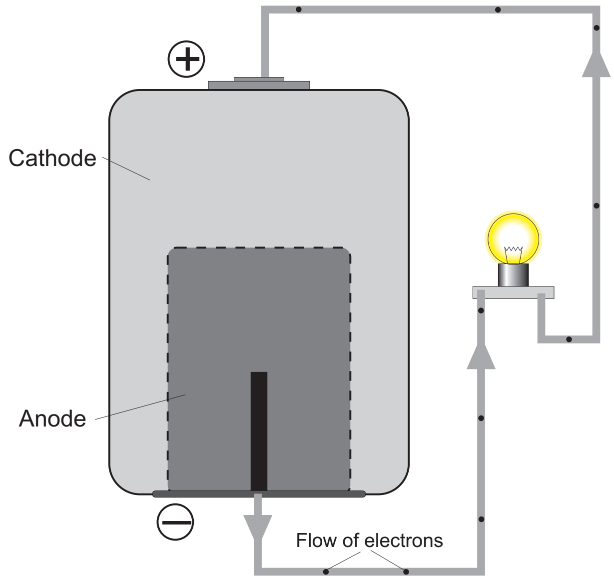


Figure 6.2: A dry cell

6.1.5 Currents

If a battery is connected by a piece of wire, the potential difference between the two ends causes a flow of electrons: The chemical energy stored in the battery is converted to the energy of the electrons. If one Coulomb of charge flows past any point in one second, then one says that a **current** of one **Ampere**, (A), is flowing. This unit is named after a French physicist who studied the forces between current carrying conductors.

Wires made of different materials respond in different ways to a potential difference applied to their ends. The German physicist Ohm discovered an empirical relation obeyed by most materials,

$$V = IR\,, \tag{6.4}$$

where V is the potential difference, I the current flow and R a constant of proportionality called the **resistance** of the material, which depends on the type of material, its size and temperature. The above law is known as **Ohm's law** and the unit of resistance Volt/Ampere is known as the **Ohm**, (Ω).

A word about conventions: For historical reasons the direction of current drawn in diagrams is that which a positive charge would take if it was the carrier: This is called the **conventional current** flow. The actual current due to the motion of electrons is opposite to the conventional current!

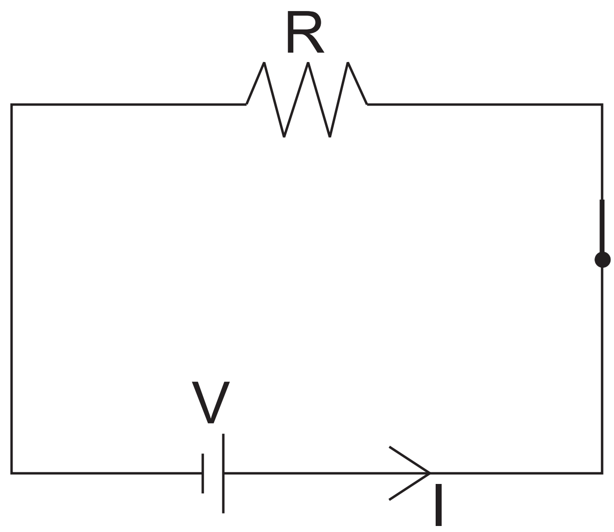


Figure 6.3: A Simple Electric Circuit

Materials such as plastic and wood are poor *electrical conductors* of electricity and are termed *electrical insulators*. A measure of resistance which is independent of the size of the wire is given by the **resistivity** ρ which is related to the length l, cross-sectional area A and and resistance R of the wire by

$$R = \frac{\rho l}{A}\,. \tag{6.5}$$

It is important to note that while the electric *current* or field flows near the speed of light, the *electrons* themselves flow at a much slower *drift velocity* ($\sim 0.01 cm/s$) because of collisions with atoms. Also, although there

Material	ρ ohm-meter
Conductors	
Silver	1.59×10^{-8}
Copper	1.68×10^{-8}
Mercury	98×10^{-8}
Semiconductors	
Graphite (carbon)	$(3-60) \times 10^{-5}$
Germanium	$(1-500) \times 10^{-3}$
Silicon	$0.1-60$
Insulators	
Glass	$10^{9}-10^{12}$
Rubber	$10^{13}-10^{15}$

Table 6.1: Resistivity at Temperature 20^0C

is current flow, the wire itself is electrically neutral because of the lattice of positive charged atoms!

The effect of even small current flow through the human body can be dramatic. As little as $0.015A$ might cause loss of muscle control and $0.07A$ might be fatal if it lasts for more than 1s.

6.1.6 Capacitance

We mentioned the Leyden jar earlier, a device used to store charge. When a charge, Q, is placed on an object, an electric field is created. In other words, the object attains a potential difference, V. The **capacitance**, ($\mathcal{C}$), of an object is a measure of the amount of charge it can store for a given potential difference,

$$\mathcal{C} = \frac{Q}{V} \tag{6.6}$$

The unit of capacitance is Coulomb/Volt, or **Farad**, named after Michael Faraday the inventor of the electric motor.

6.1.7 Power

We know that when current flows through a wire, the wire heats up. This is due to the collison of the moving electrons with the atoms in the wire. Thus the original chemical potential energy of the battery is first converted into kinetic energy of the electrons, and then into the vibrational energy of the atoms, which manifests itself as heat.

One can compute the rate at which energy is transferred between two points: If the potential difference across the points is V and the current flow is I, that means from our earlier definitions, that $V \times I$ Joules of energy per second are transferred.

Thus the Power, P, flow is

$$P = VI\,. \tag{6.7}$$

6.2 Magnetism

In addition to electric charges and forces, it has been known for a long time, in fact for thousands of years, that analogous but apparently different **magnetic** forces exist. Some naturally occuring minerals, although electrically neutral, are strongly attracted to iron. It was also discovered that these magnetic materials if left free to rotate, always orientated themselves with the Earth's north-south axis, and so already by 1000 A.D. the Chinese used such substances to make compasses to help them navigate.

Latter experiments by William Gilbert in the sixteenth century and others revealed that all magnetic materials always have two opposite poles: A **north pole and a south pole**, with like poles of materials repelling and unlike poles attracting. This situation is similar to that for the two signs of electric charges but the crucial difference is that **isolated magnetic poles (magnetic monopoles) do not exist in nature**. Thus unlike the case for electric charges, one cannot create magnetised substances with an excess of one magnetic pole over another.

(The non-occurence of magnetic monopoles in Nature is somewhat of a mystery for modern theoretical physics as almost all of the unified theories of fundamental forces predict their existence. Thus understanding why they are not seen might require some deep new concepts or principles. On the other hand the existing magnetic phenomena have been understood in terms of electric charge, as we see below).

Just as we introduced the electric field, we also have the concept of magnetic field, denoted by $\mathbf{B}$. The magnetic field, for example, describe how two magnets affect each other. The magnetic field lines around a magnet are shown in the figure.

6.2.1 Electric Motors

An important accidental discovery was made by Oersted in 1820: He found that electric currents produce magnetic fields ! For a straight conductor the direction of the magnetic field is given by the right-hand-rule: If the thumb points along the direction of the conventional current, then the fingers curl in the direction of the magnetic field. The same rule applied to segments

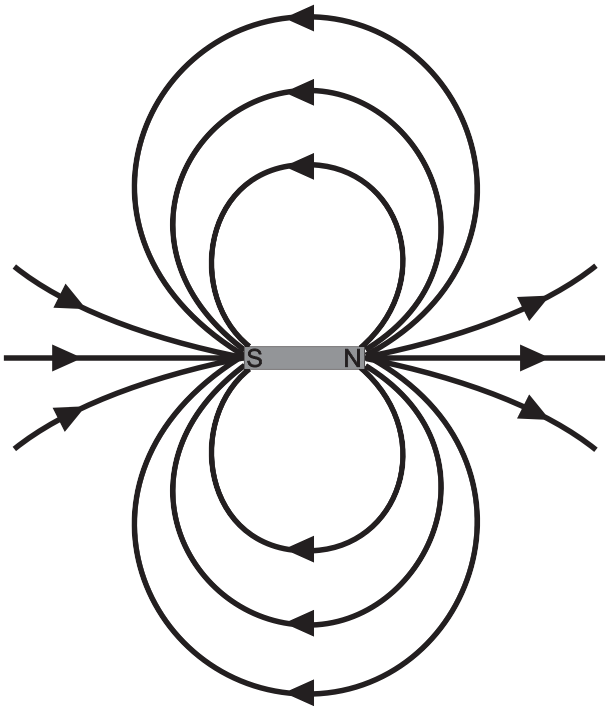


Figure 6.4: Magnetic field lines due to a magnet

of curved conductors helps one deduce the magnetic field configuration as shown in the figure.

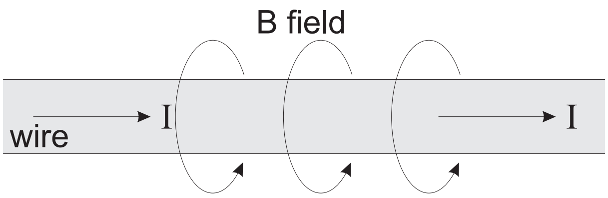


Figure 6.5: Magnetic field lines aroud a straight conductor

The creation of magnetic fields by currents is behind the operation of electromagnets which are found in numerous modern appliances. Perhaps the most important device that uses electromagnets is the **electric motor**. A DC motor, such as those used in toys, runs on **direct current** (current that does not change direction, as supplied by a battery for example). The current powers an electromagnet (the rotor) which is placed between permanent magnets. Torques cause the rotor to rotate when the poles of the electromagnet are close to poles of the same polarity on the permanent mag-

Phenomena	Magnetic Field
At Neutron Star's surface	$10^6 T$
Large Electromagnet	$1.5T$
Small Bar Magnet	$10^{-2}T$
At Earth's surface	$10^{-4}T$
Quantum flux of magnetic field	$4 \times 10^{-17} T - m^2$
Detectable magnetic field	$10^{-14}T$

Table 6.2: Some Approximate Magnetic Fields

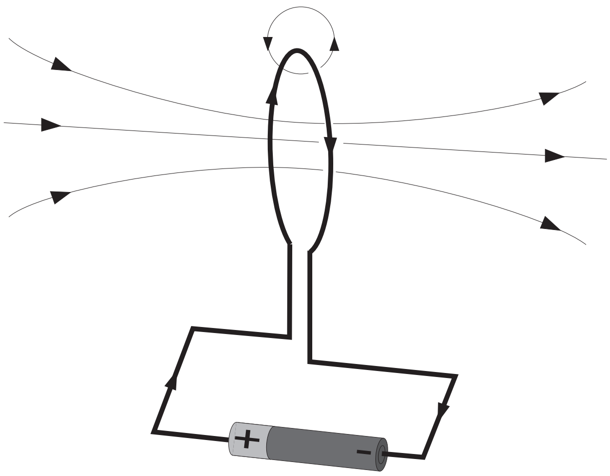


Figure 6.6: Magnetic field lines around a coil

nets. Thus to continue the rotation, the poles on the electromagnet must change polarity every half cycle. This is done by using a commutator as shown in the figure.

The well known forces between magnets implies, via Oersted's discovery, forces between one magnetic field and a carrier of electric current. Since current is just moving charge, it is no surprise then that an electric charge moving in a magnetic field experiences a force. For a charge Q moving at velocity v in a region of space with both electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$, the full force equation, in vector notation, is given by

$$\mathbf{F} = Q\mathbf{E} + Q\mathbf{v}\times\mathbf{B} \tag{6.8}$$

The equation above is called the **Lorentz equation**. Note again that this fundamental equation makes no reference to "magnetic charge" as all isolated charges are electrical and magnetic effects can be understood as due to moving charges.

From the above equation it follows that the SI unit of $\mathbf{B}$ is given by

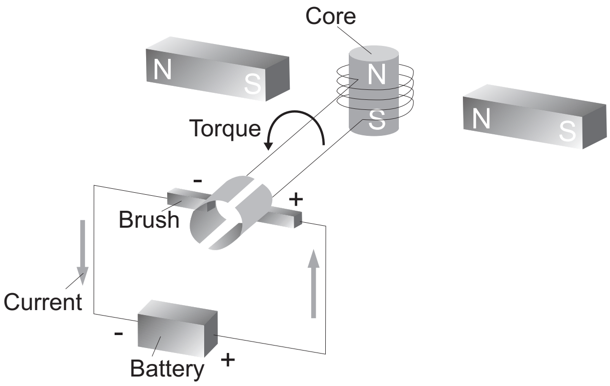


Figure 6.7: DC Motor

(Newton per (coulomb-meter/sec) or **Tesla**, named after the engineer and inventor who designed the power station at Niagara Falls.

6.2.2 Electrical Generators

Michael Faraday discovered an effect which is the converse of that found by Oersted: A changing magnetic field produces electric fields! This process of *electromagnetic induction*. If the magnetic flux through a coil is changed, the induced electric fields produce currents which tend to oppose the change (Lenz's law).

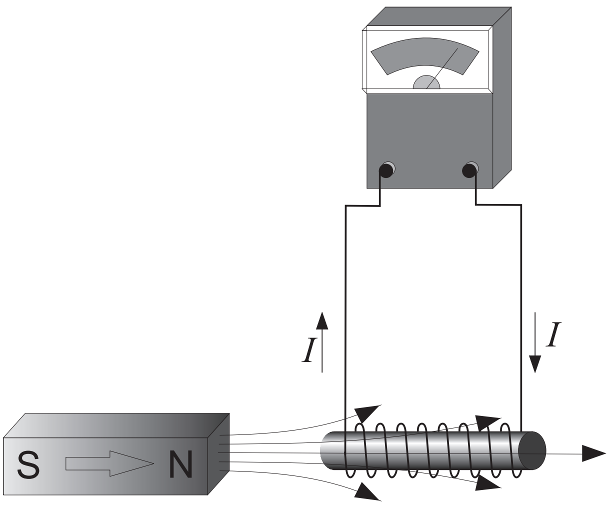


Figure 6.8: Current induced by a changing magnetic field

Electromagnetic induction is responsible for the functioning of electrical

generators which supply us with electric power. A simple generator has a rotating permanent magnet between two coils. As the magnet rotates, the flux through the coils changes, inducing an electric current through the coils. However as the magnet rotates, it presents alternate poles to each coil every half-cycle, and so the direction of the induced current also changes. This kind of current is called an **alternating current**. $50 - 60$ Hz are the common frequencies used.

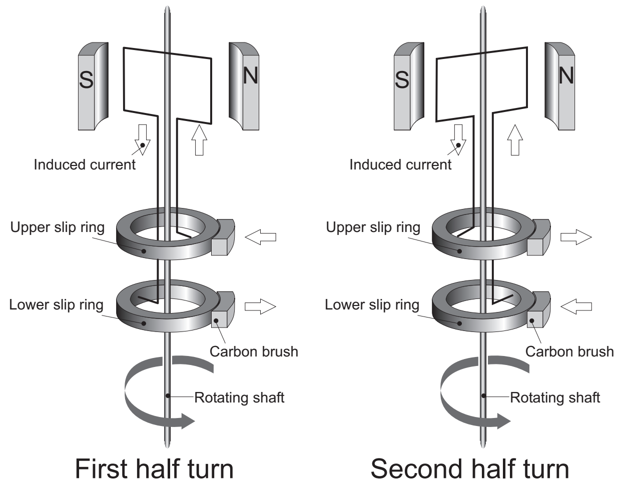


Figure 6.9: Electric Generator

Although the current is changing direction many times a second, this does not affect many devices such as light bulbs. Suppose an AC voltage $V(t) = A\sin(\omega t)$ is applied across a resistance R. The instantaneous thermal loss of power is then $V^2(t)/R = A^2\sin^2(\omega t)/R$ which is always positive. The average of that over one cycle is given by $P_{av} = A^2/2R$. This is the same power supplied by a DC voltage of amplitude $D = A/\sqrt{2}$, and quoted AC voltages refer to the equivalent DC voltage. So, 240 Ac voltage actually means an equivalent 240 DC voltage power source.

Power plant generators are usually driven by turbines that are turned by steam.

Note that the same arrangement of coils and permanent magnets as above is what is used in a synchronous AC motor, but now instead of turning the magnet by hand, one passes an AC current through the coils causing the rotor to turn !

6.3 Electric Power Distribution

The first attempt at commercial electricity distribution was by Edison in New York City in 1882. Edison supplied direct current to his customers.

Since the loss of power due to the resistance, R, of the supplying cables is I^2R when the current supplied is I, therefore for a fixed power $P = IV$ supplied at a potential difference V, the loss is equal to P^2R/V^2. Minimising the losses implies using low resistance cables (good conductor such as copper, and thick and short if possible), and/or providing the power at very high voltages (that is, low current). In order to supply the power at large distances, economics forces us to use very large voltages. However large voltages are not practical (nor safe) for common household applications and so DC power is not a feasible solution for commercial electricity distribution.

Alternating current has the advantage that it can be transmitted over long distances at very high voltages and then *stepped-down* using a **transformer** to lower voltages, with little loss. Transformers make use of the inter-relationships between electric and magnetic fields. As shown in the figure, the changing current in the primary coil produces a channging magnetic field which induces a current in the secondary coil. The voltage induced in the secondary coil (V_2) is related to the primary voltage V_1 through the relation

$$\frac{V_2}{V_1} = \frac{N_2}{N_1} \tag{6.9}$$

where N_1 and N_2 are the number of turns in the primary and secondary coils. Since energy is conserved however, the respective currents satsify $I_1V_1 = I_2V_2$.

Thus at the power station, typically some fuel is burned and the thermal energy is used to perform mechanical work that drives a generator to produce electricity. A step-up transformer raises the voltage to a few hundred thousand volts for long-distance transmission. At the destination, a step-down transformer reduces this to a smaller voltage which is further reduced by another step-down transformer for home use.

6.4 Exercises

1. Try the following experiment in an environment with low humidity: Comb your hair for some time and then bring the comb close to some small pieces of paper. Describe what you see and explain the results. Why does this experiment not work if there is considerable humidity?

2. Explain, with the help of examples, the difference between an (i) electrically neutral object, (ii) an electrically charged object and (iii) and electrically polarised object.

3. (a) Does the North end of a compass needle point exactly to the geographical North pole ?
   (b) Where is the true magnetic north pole of the Earths magnetic field

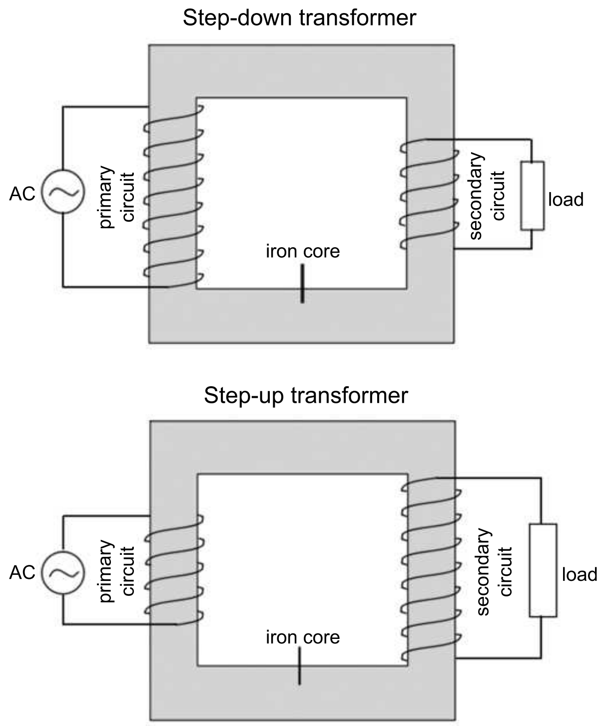


Figure 6.10: Transformers

?
(c)Do animals use the Earth's magnetic field for navigation? How ?

4. (a)Compare the magnitude of the gravitational force between the electron and proton in a hydrogen atom with its electrostatic force.
(b)If electrical forces are so large why dont we sense them (relative to the gravitational force) in daily life ?
(c) The electrons in an atom are at least 10^5 times away from the atom than the nucleus. Thus the atom is quite "hollow". Matter is then mostly empty space. So what is it that gives matter its "solid" feel?

5. (a) How many electrons would it take to form one Coulomb of charge?
(b) If the starter motor of a car draws a current of 200 A when connected to the 12V battery, how much power does it consume ?

6. (a) Check the power rating of three electrical appliances in your home and compute the current that they draw at full power.
(b) Assuming the loads to be pure resistance, compute the resistance in each case.

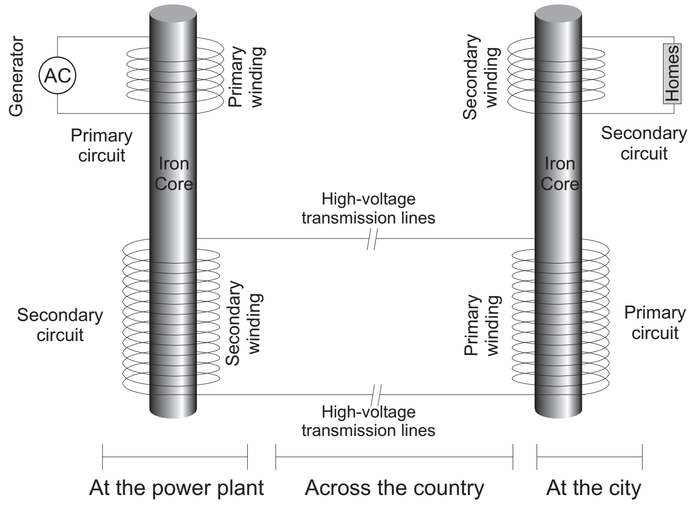


Figure 6.11: Electric Power Distribution

7. (a) In a flashlight, two 1.5V batteries power a light bulb of resistance 2 Ω. How much current is flowing in the circuit? How much power is consumed?
(b) What is the resistance of a household 60 W lightbulb if it draws a current of 1 A ?
(c) A car battery is recharged by running a current of 4 A at 12 V for one hour. How much energy was used to recharge the battery ?

8. (a) The resistance of the human body ranges from $500,000\Omega$ when dry to about 100Ω when wet with salty water. Compare the current flows through the body if the two hands are used to complete an electrical circuit between a 12V battery.
(b) Explain then why one sees warning signs of "Danger – High Voltage", rather than "Danger – High Current".
(c) Why is it that birds can perch on high-voltage lines without getting a shock ?

9. (a) If an extension cord is overloaded so that it carries too much current, it overheats. Why ?
(b) How does a fuse limit current flowing through an electric circuit ?
(c) Why does a short-circuit in an electrical appliance cause the fuse to melt ?
(d) What is are the roles of the three-pins of an electric plug ?

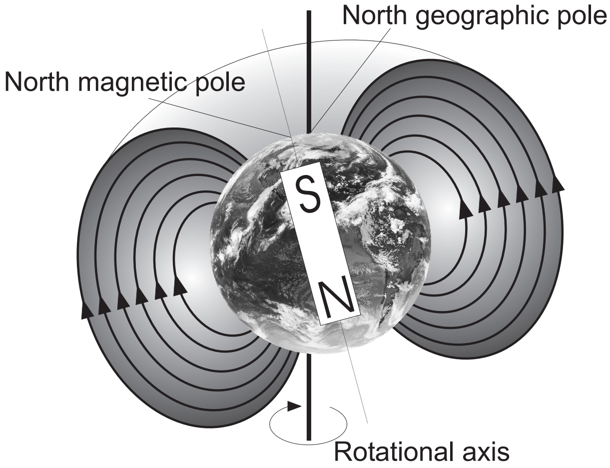


Figure 6.12: Earth's magnetic field

10. What factors influence the choice of material for use in electrical cables? Can you use those factors to deduce the best material ?

11. Check your latest home electricity consumption bill (or estimate it using Ref.[4]).
(a) What units is the energy consumption quoted in? How is that related to Joules?
(b) What is the rate charged for the energy use?
(c) Find out which tools or appliances in your home are the top three energy consumers.
(d) Estimate how much energy is wasted in your home each day.
(e) What are the consequences of wasted energy?
(f) How can you be more efficient and environmentally friendly?
(g) Read the tips for home energy conservation at Ref.[5] and discuss why they are sensible.

12. (a) How high would you have to walk up to have a gain in potential energy equivalent to your average daily home energy consumption bill?
(b) What is the mass equivalent of that energy consumption?

13. (a) What is the difference between "Direct Current" (DC) and "Alternating Current" (AC) ?
(b) Why is alternating current commonly used for transmission of power ?
(c) At what frequency is the AC power transmitted ?
(d) How may DC power be produced? (Think of as many ways as you can. I have four !)

14. (a) How would a magnetically levitated (MAGLEV) train operate ? (b) What would be the advantages and disadvantages of such trains compared to conventional trains that use wheels? (Hint: See Ref.[6].) (c) How do MAGLEV trains "brake" to come to a stop ? (d) How do maglev trains compare with those that might use the Wind-in-Ground effect Ref.[7] ?

15. Explain, with the help of diagrams, the workings of universal and induction motors. (Hint see 1).

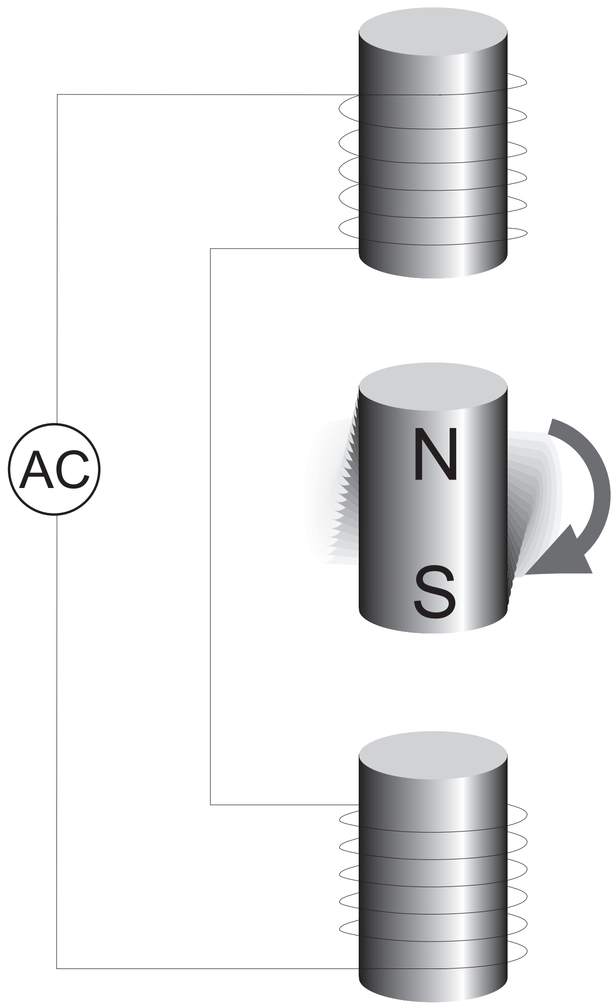


Figure 6.13: AC Motor

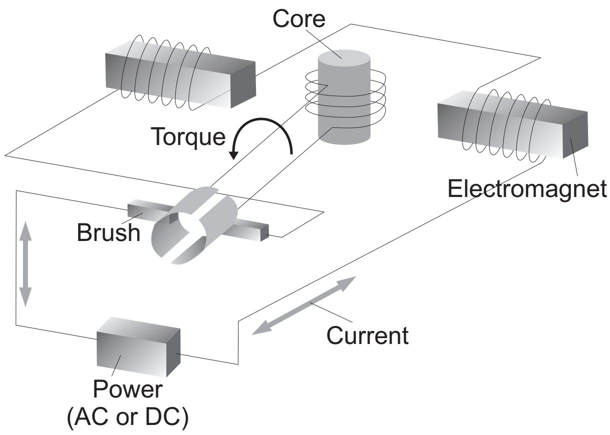


Figure 6.14: Universal Motor

16. Make lists of all appliances that you encounter in your daily life that use the following:
(a) A DC electric motor
(b) A synchronous AC motor
(c) A universal motor
(d) An induction AC motor
.

17. (a) How does an audio speaker use magnets and electricity to convert electrical signals into sound ?
(b) What are teh basic operating principles of a simple microphone?

18. (a) Can a transformer be used to step-up or step-down DC voltage ? Why or why not ?
(b) If the voltage carried by a transmission line was doubled while keeping the power delivered constant, by how much would the power loss be reduced ?

19. (a) What are the different forms of energy that electric power plants use to generate electrical power?
(b) Which of those is the most efficient from the energy conversion point of view ?
(c) Which is the least polluting ?

20. Optional: (a) Can plastics conduct electricity ?
(b) What is a semiconductor and what use does it have ?
(c) What is a superconductor and what use does it have ?
(d) Why are good conductors of heat also good conductors of electricity ?

6.5 References

1. The Electric Motor at http://www.howstuffworks.com/motor.htm and also a downloadable pdf file from Wiley How Things Work website at http://www.lenoxsoftworks.com/htw/ch09/ch9_5/electric_motors.pdf

2. Electricity Generation Downloadable pdf file from Wiley How Things Work website at http://www.lenoxsoftworks.com/htw/ch09/ch9_4/electric_power_generation.pdf

3. The telephone at http://www.howstuffworks.com/telephone.htm

4. Consumption calculator at http://www.psupply.com.sg/cs_services_energy-audit.asp

5. Home Energy Conservation Tips at http://www.ema.gov.sg/Electricity/efficiency.php and http://www.psupply.com.sg/cs_EEC_frameset.asp?ID=

6. Maglev Trains at http://www.howstuffworks.com/maglev-train.htm

7. Flying trains at http://www.abc.net.au/science/news/stories/s108912.htm

Chapter 7

Electrodynamics: From Radio to Light

OK, so what's the speed of dark?
–Steven Wright

Before James Clerk Maxwell involved himself, the study of electric and magnetic phenomena was dominated by a number of distinct laws, such as Coulomb's law, Faraday's and Ampere's laws and others. Maxwell noticed that all the electric and magnetic laws of the time could be summarised by four equations. However he also noticed that the four equations, taken together, were incompatible with the principle of charge conservation.

Now, charge conservation was considered a physically reasonable thing and so difficult to ignore. Hence Maxwell decided that one of those four laws had to be modified if charge conservation was to hold, and as a consequence he found that the modified equations, **predicted** the existence of a new phenomenon: *Electromagnetic waves* which travelled at the same speed as the then known speed of light! Maxwell then hypothesized that light itself was just one form of the elctromagnetic wave. Maxwell's predictions were experimentally verified by Hertz and others and that gave birth to such new areas of activity as radio communication.

The fundamental laws of electricity and magnetism as summarised by Maxwell are called **Maxwell's equations**, and describe the unified phenomenon of **electromagnetism or electrodynamics**. That is, electric and magnetic fields are really the same entity, and appear to be different only in certain special circumstances. A charge moving with a constant velocity, for example, generates a combination of both electric and magnetic fields (recall that an electric current produces a magnetic field and that current is simply electric charge in motion).

Maxwell's equations are so beautiful (they have more symmetry than Newton's equations) that they helped inspire Einstein in formulating his

Special Theory of Relativity. (Yes, mathematicians, physicists and engineers do talk of beauty too – but sometimes use it also for equations and the such).

In summary, the partial unification of electricity and magnetism by Faraday and others was completed by Maxwell who then went on to achieve another unification: That of electricity, magnetism and light !

7.1 Electromagnetic Waves

Recall that a collection of moving charges generates electric and magnetic fields **E** and **B**. The $\mathbf{E}(t,x)$ and **B**(t,x) are vectors that depend on time, and exist over all space; that is, at every point x in space and at every instant t, three real numbers specify the electric field, and similarly for the magnetic field.

Maxwell deduced that oscillations of the electromagnetic field gives rise to **electromagnetic (EM) waves**. In contrast to sound waves, for EM waves there is **no underlying material medium** whose oscillations result in the wave phenomenon: That is, EM waves can propagate in a vacuum! Another important difference between EM waves and sound is that **EM waves are transverse waves**, in contrast to sound which is a longitudinal wave. That is, the wave oscillations for EM waves are perpendicular to the direction of propagation of the wave, wheras for sound waves the oscillations are in the same direction as the propagation of the wave.

EM waves are said to be **linearly polarised**, that is, the plane in which the **E** oscillates is fixed, **circularly polarised** (when the plane of polarisation rotates), or **unpolarised** when the **E** field changes direction randomly.

All electromagnetic waves travel at the same speed through vacuum: At the speed of light, which is about 3×10^8 *m/s.*

Wave Equation
In one dimension, x, the equation describing wave propagation is of the form

$$\frac{\partial^2 f(x,t)}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f(x,t)}{\partial t^2} = 0\,, \tag{7.1}$$

where $f(x,t)$ is the changing quantity charactersing the wave (e.g. pressure amplitude for sound or electric field for EM waves) and t is the time variable. It is easy to check by substitution that

$$f(x,t) = A \sin 2\pi(\frac{x}{\lambda} - \frac{t}{T})\,, \tag{7.2}$$

is a solution of the wave equation if $v = \lambda/T \equiv \lambda f$ (the constant A is the amplitude of teh wave). From the form of the solution one sees that λ is the wavelength of the wave, that is, the distance between two points on the wave that are in phase, while T is the period or time taken for the wave at a fixed loaction to repeat its value. The frequency $f = 1/T$ is therefore the number of oscillations of the wave in one second and

$$v = f\lambda\,, \tag{7.3}$$

is the wave speed.

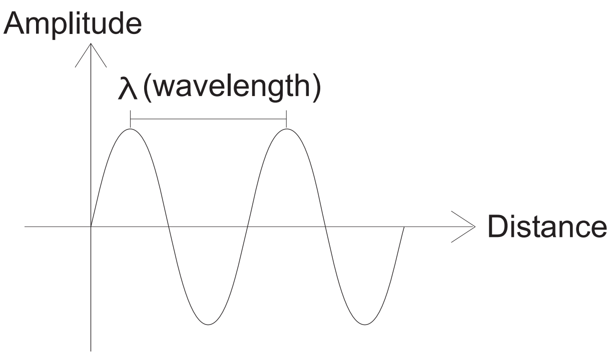


Figure 7.1: Definition of Wavelength

7.1.1 Generation and Propagation of EM Waves

An **accelerating** elecric charge radiates electromagnetic waves, which are just oscillations of the elctromagnetic field.

If the electric charge is made to oscillate at a frequency f, it will produce electromagnetic waves of frequency at the same frequency. In an antenna charge is made to oscillate along the length of an antenna, resulting in the

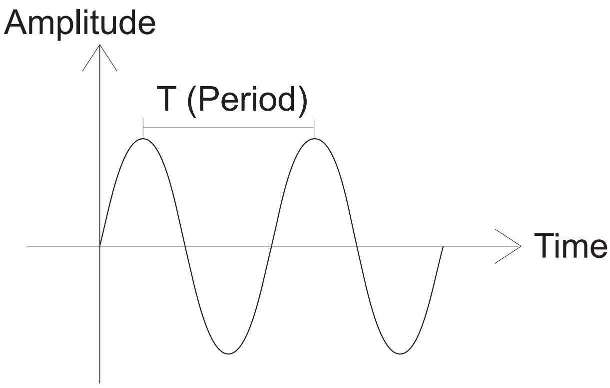


Figure 7.2: Definition of Period

emission of electromagnetic radiation such as radio waves. The waves are sustained by the work done to accelerate the charge.

How does EM radiation propagate in empty space ? This is possible because a changing electric field creates a magnetic field, and in turn the changing magnetic field creates an another electric field. It is this mutual feedback that sustains the propagation of an electromagnetic wave.

Let us consider for simplicity electric and magnetic fields far from the charges that have generated the radiation. If the electromagnetic wave is propagating in vacuum along the x-direction, then since it is a transverse wave, the oscillations of the electric and magnetic field are in the directions perpendicular to the direction of propagation, and hence lie in the yz-plane. For example, in a linearly polarised wave, if the electric field **E** lies along the y-axis, then the magnetic field **B** lies along the z-axis.

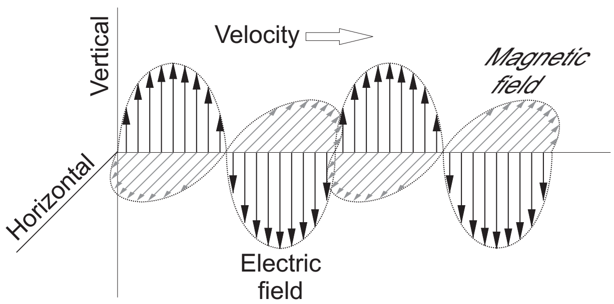


Figure 7.3: Electromagnetic Wave Propagation

Energy is propagated along the direction of motion of the wave, which in this example is in the x-direction. The **intensity** of the wave is a measure of how much radiant energy per meter squared is received at a point each

second, and is proportional to $\mathbf{E}^2$. At a large distance R from a small source, the intensity decreases as $1/R^2$, explaining why in many applications transmitters have to be very powerful.

Electromagnetic Shielding
Consider a solid conductor. If it is placed in a static external electric field, the electric fields will drive the electrons in the conductor until elecrostatic conditions are reached whereupon *no electric fields can exist inside the conductor*: That is, the induced charges on the conductor will arrange themselves to neutralise any electric field inside the conductor. Gauss's Law then implies that there will be no induced charges inside the conductor either — they will reside only on the surface.
Remarkably it can be shown that the above results are true even if the conductor is hollow: No electric fields exist under electrostatic conditions inside the cavity, nor do any charges reside on the inner surface of the cavity.
If the external field is not static, then electric fields do penetrate slightly, by a distance known as the skin-depth.
This screening property of good conductors is made use of in blocking stray electromagnetic radiation from computers and in protecting sensitive equipment.
This is also the reason why staying inside a fully closed car during thunderstorms provides protection from lightning.

7.1.2 Names Given to EM Waves

The figure below shows the elctromagnetic spectrum and the names given to the different wavelengths. Recall that the wavelength of a wave is the distance between two points on the wave which are in phase, for example, the two peaks in the propagation of the wave. Frequency is the number of times the wave oscillates each second. So for a propagating wave, the speed (c) is related to the wavelength (λ) and frequency (f) by the relation

$$c = \lambda f \tag{7.4}$$

Light is just a special case of electromagnetic waves: Those with wavelengths in the range of 10^{-6}m to 10^{-7}m form what we call the visible spectrum. We will discuss it in more depth later.

Note: When waves propagate from one medium to another, it is their frequency that remains unchanged while their wavelength changes, so that the speed of EM waves in a medium actually depends on their frequency. For example, white light spilts into several colours when pasing through glass because the low frequency components (red) travel faster in glass (or water) than the high frequency (blue) ones.

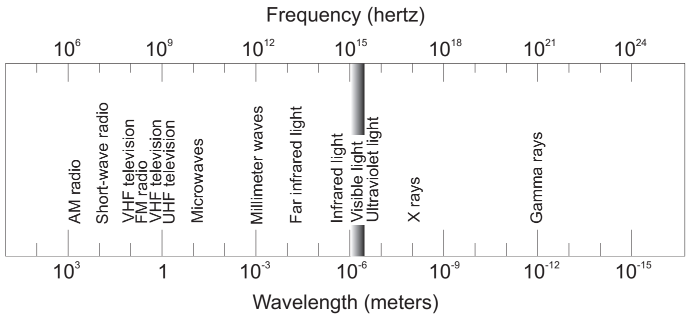


Figure 7.4: The electromagnetic spectrum

Faster than Light
The special theory of relativity forbids objects with real mass, or information, to propagate faster than c, the speed of electromagnetic waves (like light) in a vacuum.
Although EM waves do not require a medium for their propagation, they are affected by interactions with charges in any medium in which they travel. The interactions do not change the frequency of the wave but result in a wave of modified wavelength, and hence of different speed. This is why light bends towards the normal when it travels from air to glass — its speed in glass is less than its speed in air.
The speed v of light in a medium with *refractive index* n is given by

$$v = \frac{c}{n} \tag{7.5}$$

where n depends on the frequency of the wave (so that different frequencies are refracted by different amounts causing effects like the rainbow).
It is possible for particles in a medium to travel faster than c/n, *the speed of light in the medium*, while still moving slower than c! When charged particles move through a medium faster than c/n, the radiation they emit forms a cone similar to the bow wave formed by a speedboat. By measuring the angle of the cone, one can deduce the speed of the particle.
Such radiation is called **Cherenkov radiation** after its discoverer who, together with Tamm and Frank, won the Nobel prize in 1958.

7.2 Radio

As mentioned above, accelerating charge in an antenna leads to the transmission of electromagnetic waves. If the antenna is part of a radio transmitter

then it is creating a radio waves with the same frequency as the oscillations. When this radio wave encounters another antenna (the receiver), it causes charges in the reciver to oscillate with that same frequency too.

An essential component of the transmitter and receiver is the tuner, which basically consists of a capacitor and inductor (coil) in series. By adjusting the values of the capacitor and inductor, the tuner can be made to resonate at different frequencies. The resonance amplifies both the transmitting and receiving signals sufficiently to make radio practical. When one tunes to a different staition, one is basically changing the resonant frequency of the receiver so that one particular frequency is amplified and selected.

Here is a qualitative explanation of why a tuner oscillates and how resonance occurs. Suppose a charged capacitor is connected to an inductor. Then current will flow through the circuit as the capacitor discharges itself. The current creates a magnetic field which is changing and so another electric field is induced to try and oppose those changes. This opposition to quick changes means that even when the capacitor is fully discharged, the current keeps flowing in the same direction, driven by the energy now stored in the inductor. Eventually the energy in the inductor reduces to zero and the capacitor becomes fully charged again but now with an opposite polarity ! The process therefore repeats, giving rise to periodic oscillations in the current. This tuner, or ”LC” circuit, has a **natural frequency** of oscillation and if an external current is imposed onto the circuit with the same frequency then a **resonance** occurs: The resultant current has the maximum amplitude.

The mechanical anlaogy is that of a playground swing. It has a natural frequency of oscillation and if one pushes on it with that same frequency, then the amplitude of its oscillation increases.

7.2.1 FM vs AM

The purpose of radio waves is of course to transmit information about sound. A pure sinusoidal wave actually contains almost no information since it is perfectly periodic. Two ways in which the relevant information is encoded into radio waves is **frequency modulation** and **amplitude modulation**. As the name suggests, in frequency modulation, the radio waves frequency is increased or decreased slightly to represent the corresponding compressions and rarefactions of the sound wave. In amplitude modulation, it is the amplitude of the wave which is increased or decread slightly to represent the relevant deformations of the sound wave.

Commercial AM bands are between 550 kHz and 1600 kHz while FM is between 88 MHz and 108 MHz. An important consideration is the **bandwith**: The range of frequencies around the *carrier frequency*. FM radio stations have 200kHz of bandwith, with the carrier frequency in the middle. The range allows a broad range of audio frequencies to be represented al-

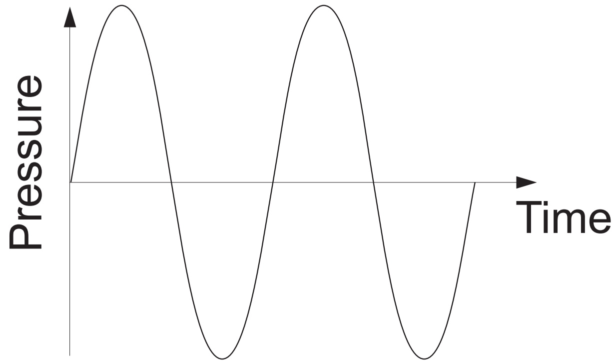


Figure 7.5: Unmodulated Wave

lowing for better music. By contrast AM radio stations are allocated only 10kHz of bandwidth, with the carrier frequency again in the middle of the range.

7.3 Television

Representing a picture using electromagnetic waves obviously is more complicated than representing sound. We will discuss here mainly the process for black and white analog TV transmission.

7.3.1 The Picture Tube

In the picture tube, electrons are accelerated from the cathode to the screen where they strike the phosphorous coating. The kinetic energy of the electrons is transferred to the phosphor which glows by emitting white light. By making the electron beam scan the screen, dots of light glow, creating an image which we perceive. Of course the individual dots must light up fast enough in succession so that we get the illusion of a one continuous picture. Indeed the beam starts at the upper left hand corner of the screen and moves to the right. Once a line the beam can move to the next line and so on until it reaches the bottom of the screen. Such a procedure would take 1/30th of a second which is a bit slow as our eyes will perceive the flicker. The trick which is actually adopted is to first scan all the odd number lines and then go back to the top to scan the even numbered lines. In this way a picture frame is formed every 1/60th of a second which is sufficient for smooth viewing.

Electric and magnetic fields within the tube are used for focusing the beam. The horizontal and vertical deflection of the electron beam is performed by magnets using the Lorentz force discussed earlier: A vertical

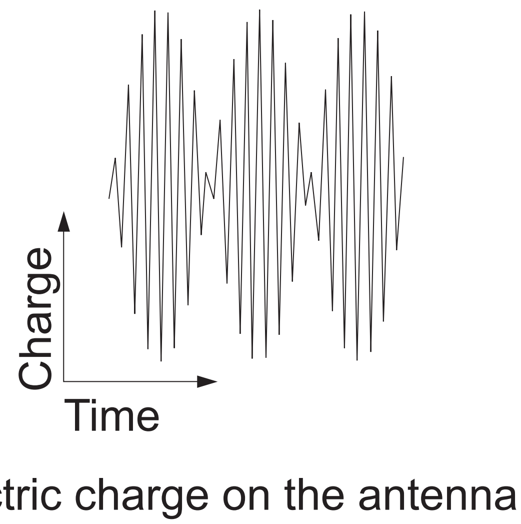


Figure 7.6: Amplitude Modulated Wave

magnetic field deflects the electrons horizontally while a horizontal magnetic field deflects the electrons vertically. Shades of gray on the screen are produced by controlling the amount of current in the beam. This is done by changing the amount of negative charge on the grid in front of the cathode: Thus effectively electrons are repelled by the grid and so not all of them make it to the anode and then on to the screen.

A colour picture tube has three separate electron guns and a screen coated with phosphor dots of the three primary colours: Red, Green and Blue. Between the beams and the screen are three masks with holes aligned so that each beam can only strike dots of one colour.

7.3.2 The EM signal

The black and white video signal is basically as shown in the figure. It represents the brightness along each horizontal line, from white to black, with a small jump beyond black indicating a jump to the next line. The beginning of a new picture, that is the movement of the beam back to the top left hand corner, is indicated by a bigger jump beyond black.

That is, the video signal is amplitude modulated but because of the need to represent more information every second, video signals require much more bandwidth than radio signals. For example the NTSC standard has 6MHz of bandwidth. What about the sound part of the television signal ? For analog TV this part of the signal is frequency modulated.

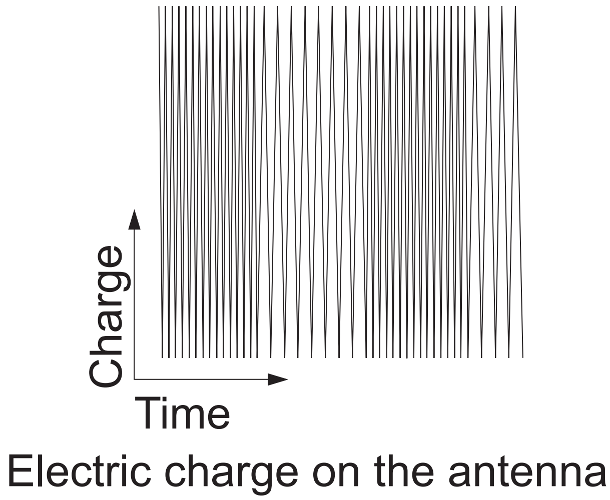


Figure 7.7: Frequency Modulated Wave

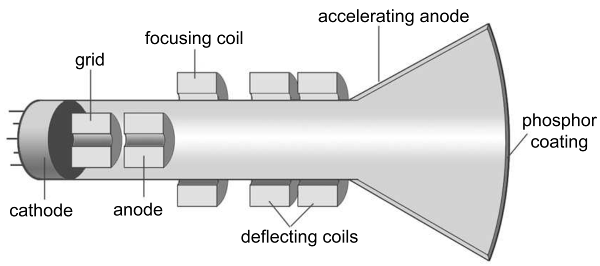


Figure 7.8: Cathode Ray Tube

7.4 Cable Transmission

Instead of transmitting EM waves through open space, one may also send them along cables. A coaxial cable consists of along the axis of a metal tube. This arrangement allows EM waves to propagate along the cable at high speeds (less than the speed of light), even when the cable is not straight. One of the advantages of cable is that since the waves inside cannot interact with those outside, one can use more of the EM spectrum for TV channels.

7.5 Microwave Ovens

The part of the EM spectrum ranging from 1m to 1mm is referrred to as microwaves. A microwave oven uses EM waves of 2.45 GHz (12.2 cm) to cause a forced oscillation of water molecules in foods. Recall that water is a polar molecule because of the oxygen atoms greater attraction of electrons

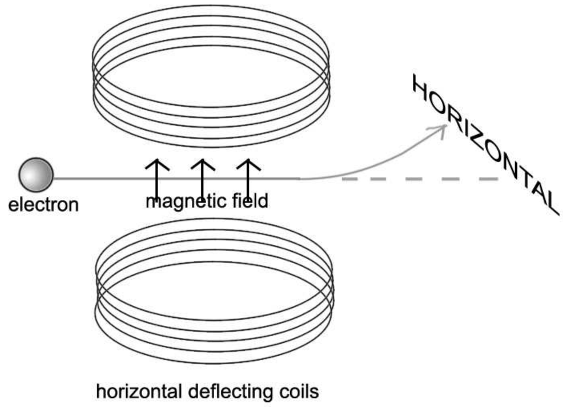


Figure 7.9: Horizontal deflection of elctron beam

in the covalent bond it forms with the hydrogen atoms. That is, the water molecule is like an electric dipole which responds to a fluctuating electric field by oscillating.

However as the molecules rotate, they collide with other nearby molecules and thus some of the energy is converted to thermal energy ("random" agitation of the molecules). This is how cooking takes place in a microwave.

The walls of the microwave are metallic and so reflect the microwaves. The metal grid on the front window also relects the waves because their size is much smaller than the wavelength of the microwave.

7.6 Light

7.6.1 Sunlight

The sunlight that we experience on the surface of the Earth is only a small portion of the solar radiation incident on the Earth, a large portion of the injurious components of that radiation being diminished by our atmosphere, by the processes of reflection, scattering and absorption.

The stratosphere (upper atmosphere) has a layer of ozone which is a strong absorber of the biologically harmful short wavelength ultraviolet radiation below 280nm. It is for this reason that there is a great deal of concern at the rate at which pollutants such as flurocarbons deplete the ozone layer allowing the ultraviolet radiation to penetrate.

The troposphere (lower atmosphere) also plays a useful role in diminishing harmful radiation. At this level, clouds (which are nothing but moisture and ice particles), and suspended matter such as dust and smoke, reflect and scatter radiation. Indeed the blue colour of the sky(and the reddish

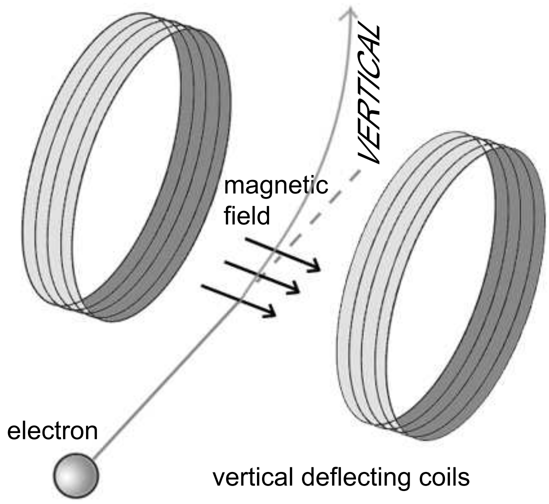


Figure 7.10: Vertical deflection of electron beam

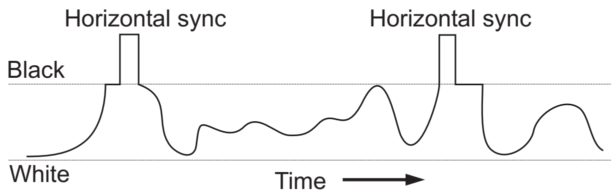


Figure 7.11: Video Signal

colour of sunrises dnd sunsets) is due to the preferential scattering of short wavelengths.

One can ask about the energy carried by each particular wavelength that comprises sunlight. This is shown in the figure. Note how the peak in the energy distribution is in the middle of what we humans call the visible spectrum. Thus light is simply that part of the electromagnetic spectrum, those with wavelength in the range of 10^{-6}m to 10^{-7}m that our eyes have become specially adapted to through evolution.

Many of you might have heard of the expression "red hot", or seen in films the yellowish colour of molten steel. *In fact the hotter an object is, the more it radiates at shorter wavelengths.* Quantitatively, the peak of the energy distribution for an ideal emitter (a "black body") is located at λ_m where

$$\lambda_m T \approx 0.003\, m\ K\,, \tag{7.6}$$

which is known as Wein's displacement law. So "blue hot" would be very

hot indeed! The sun has a surface temperature of about 5800 Kelvins, which make it somewhat "white hot".

7.6.2 Simulating Sunlight: The Light Bulb

Since we have become acustomed to sunlight, we would naturally like to produce artificial light which is similar to that from the Sun. A major technological innovation was the invention of the electric light bulb (no, it was not Edison who was first) which, because of its convenience and cleaniness, rapidly replaced candles and and lighted gas as a form of illumination both indoors and outdoors.

The simple incandescent light-bulb shown in the figure works on the principle that when an electric current passes through a conductor, heat is generated. If the conductor becomes sufficiently hot (say, more than $400°C$, then it emits light. However to get light of similar quality as that from the sun, the filament of the incandescent bulb must be heated to a temperature clsoe to $5800°C$: At $500°C$ we would only get dull red light, while a candle flame gives its usual orangish light at $1700°C$.

Unfortunately we do not know of any substance that remains solid up to $5800°C$, so the best that can be achieved is to use tungsten which can be heated to about $2500°C$, giving off yellow-white light. To maximise the heat generation a long, thin piece of tungsten is used. A piece of wire about $0.5m$ long is fit into a small bulb by winding it in the form of a double spiral.

To prevent the tungsten from burning at those high-temperatures, it is encased in a glass bulb. The inside is not a vacuum but rather filled with some inert gas such as argon or nitrogen. the function of the filler gases is to provide "back pressure" and thus reduce the evaporation of tungsten atoms from the filament. Nevertheless, the evaporation proceeds at a lesser pace causing an eventual breaking of the filament.

In addition to producing light which is more yellowish compared to sunlight, the simple incandescent bulb has another serious short-coming: At that temperature most of the energy (about 88%!) it emits is actually in the invisble infrared. This not only makes it highly inefficient but in addition the infrared radiation heats up the room unnecessarily.

7.7 Exercises

1. (a) Can you receive a radio transmission from over the horizon ?
   (b) Why are some radio stations clearer at night than in the day ?
   (c) Why is it that if one is too far from the trasmiter, then only the loud parts of an AM transmission can be heard while for FM signals it is a case of all or nothing ?
   (d) What are the other advantages and disadvantages of FM radio transmission over AM ?

2. (a) Why are commercial radio AM and FM frequencies limited to the ranges stated in the notes ? (Hint: See Ref.[1]).
(b) Make a list of the different applications that use EM waves to which you are exposed to in your daily life.

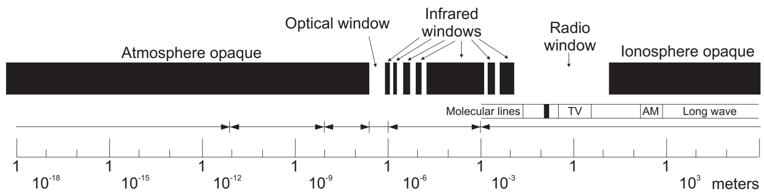


Figure 7.12: The electromagnetic spectrum

3. In our discussion of the tuner circuit, we ignored resistances such as those in the connecting wires. What effect would this have on the current oscillations ?

4. (a) Why is it that cordless phones do not function effectively when they are far from the base unit ?
(b) Why are computer casings metallic ?

5. (a) If a photograph is taken of a TV screen using an exposure time of 1/100th of a second, what would you see in the developed photo?
(b) What happens to all the electrons which strike the TV screen ?
(c) What are the advantages of cable TV compared to broadcast TV ?

6. (a) Would food cook better in a microwave oven if covered in aluminium foil ?
(b) Why are some plates and cups not "microwave safe" ?
(c) Why do foods not turn brown when cooked in a microwave oven?
(d) What safety precautions must you take when cooking and eating food cooked in a microwave oven ? Why ?
(e) Why is the door of a microwave oven designed to switch off the oven it is opened ?
(f) What are the advantages and disadvantages of microwave cooking compared to traditional forms ?

7. How does cellular telecommunications work? Read the article in Ref.[4] and answer the following questions.
(a) What does the term "cell" refer to ? How big is it ?
(b) What are the advantages of usings cells compared to older central antenna systems ?
(c) How is interference between neighbouring cells avoided ?

(d) How are calls directed to your cell phone ?
(e) What happens if while using a cell phone you cross cells ?
(f) Why is a cell phone called a duplex device ? What is an example of a simplex device ?
(g) what are some of the disadvantages of cell phones ?
(h) Can your location be tracked if you use a cell phone ?

8. The Global Positioning System or "GPS" was originally developed for military use but has been made available to the general public. With a GPS receiver a person can determine her/his exact location at any point on the Earth. Read the article at Ref.[5] and then answer the following questions:
(a) How many satellites make up the GPS system?
(b) How many are needed for your receiver to determine its exact location? Why?
(c) Are the GPS satellites in geostationary orbits?
(d) How accurate is a GPS receiver ?
(e) Why do you need an accurate world map
(f) List four different areas where the GPS system is used.
(g) Speculate on some future uses of the GPS system and the consequences for human activity that use will have.

9. (a) Can you estimate the temperature of the glowing charcoal in your barbeque pit ?
Estimate the colours of three objects at the following temperatures: $1700, 2500, 5800$ degress Celcius.

10. What advantages do halogen light bulbs have compared to normal light bulbs ? How do they work ?

11. (a) What advantage and/or protection can a good pair of sunglasses provide ?
(b)What are some of the dangers of using a cheap pair of sunglasses?
(c) How do Polaroid sunglasses work?
(d) Can one produce "Polaroid" sound filters? Why not?
(e) Try out the simulations in ref.[7] and link with what you have learned.

12. (a) Why is UV light capable of producing skin damage but visible light not?
(b) Why do some people apply sunscreen lotion when they are out in strong sunlight for long periods?
(c) Why is it that wearing sunglasses while at the beach might increase someones chances of getting sunburn ? Why is sunburnt skin reddish ?
(d) Why are some drugs packaged in amber-coloured containers ?

(e) So is there any use of exposure to sunlight ?
(Hint : see Ref.[10]).

13. Optional:
(a)If thunder is heard two seconds after lightning is seen, how far away did the event originate?
(b)Device a rule-of-thumb to determine the approximate location of a lightning strike (You may take the speed of sound to be 340m/s).
(c) Why do trees often explode when struck by lightning ?

14. Optional: Adaptive optics is a method of getting better resolution of stars by cancelling the effects of the atmosphere. Read about the latest exciting developments in Ref.[8].

7.8 References

1. The Radio Spectrum at
http://www.howstuffworks.com/radio-spectrum.htm

2. Television at http://www.howstuffworks.com/tv.htm

3. Microwave ovens at http://www.howstuffworks.com/microwave.htm

4. Cell phones at http://www.howstuffworks.com/cell-phone.htm

5. GPS at http://www.howstuffworks.com/gps.htm

6. Solar Radiation at
http://www.eren.doe.gov/consumerinfo/refbriefs/v138.html

7. Polarisation at
http://www.netzmedien.de/software/download/java/polarisation/index.html

8. Adaptive Optics at
http://www.mtwilson.edu/Science/AdapOpt/Overview/ and
http://www.aoainc.com/technologies/adaptiveandmicrooptics/aostutorial.html
and
http://op.ph.ic.ac.uk/ao/overview.html

9. Sunglasses at http://www.howstuffworks.com/sunglass.htm

10. Sunscreens at http://www.howstuffworks.com/sunscreen.htm

Chapter 8

Quantum Technologies

When you reach the end of what you should know, you will be at the beginning of what you should sense. –Kahlil Gibran

Newton's Laws are only an approximate description of our world. At the start of the last century, as matter was probed at small distances, it was discovered that Newtonian mechanics could give neither a complete nor consistent description of observed phenomena such as the discrete nature of spectral lines.

It took many years before a new consistent theory of matter emerged. This new theory is called **Quantum Mechanics**. So if everything obeys the laws of quantum mechanics, is not every technology quantum? If we are concerned only with the properties of matter at large scales, then Newtonian mechanics suffices as a very good description, as we have seen in previous chapters. However there are some technologies which can only be understood in terms of the newer (and stranger) quantum mechanics: Those that deal directly with the quantum nature of matter and energy at very small distances.

The transistor is an example of a quantum technology that has had a huge impact on society. In this chapter we will first review some basic differences between quantum and Newtonian mechanics and then highlight some "old" quantum technologies such as the laser and some emerging ones.

8.1 Origins of Quantum Theory

The initial struggle to explain the microscopic world involved a mix of classical concepts and new rules: These had success in explaining the such things as the spectrum of black-body radiation, the photoelectric effect and the stability of atoms. All these results would later follow from a self-consistent formulation of a new theory, *quantum mechanics* which we discuss briefly later.

8.1.1 Particle Nature of Light: Planck's Quantum Hypothesis

Near the end of the nineteenth century the electromagnetic wave theory of light encountered severe problems in trying to explain the observed spectrum of radiation from hot objects — "black body radiation".

Several attempts at reconciling theory with experiment failed until Max Planck proposed in 1900 that vibrating atoms that emit light do so in discrete amounts. That is, energy was postulated to be **quantised**, given by the relation

$$E = Nhf \tag{8.1}$$

where f is the frequency of vibration, N an integer, and the constant h is an empirical constant of nature, required by dimensional analysis, and is called Planck's constant. Its numerical value is given by

$$h = 6.62618 \times 10^{-34} Js \tag{8.2}$$

In classical physics, high frequency radiation such as ultra-violet was predicted to make large contribution to the black body spectrum, in contradiction to the empirical data. Planck's quantum postulate solved the problem because to emit even a single *quantum* of ultra-violet radiation would require a minimum energy much larger than the typical available thermal energy of order kT: Hence very high frequencies would not be present in the radiated spectrum of a black-body. The black-body spectrum obtained from Planck's postulate fits the experimental data perfectly.

Planck was actually hesitant to assert that light itself *propagated* in discrete packets: He considered his solution to be only mathematical. It was Einstein in 1905 who boldy proposed that light actually propagated in discrete energy packets called **photons**, with the energy of each photon being related to the wave frequency by Planck's relation. Indeed it is for his explanation of the photoelectric effect using the photon concept that Einstein was awarded the Nobel Prize in 1921. The photoelectric effect is the basic principle behind the operation of solar cells.

Note how the discrete particle-like nature of photons nevertheless depends on the frequency of a wave: **This is already a hint of a wave-particle duality to be discussed below!** *The phenomenon of electromagnetic radiation is a classical approximation to the quantum theory of photons.*

8.1.2 Stability of Atoms

At the end of the nineteenth century it was thought that the atom was a microscopic version of the solar system: A positively charged nucleus with negatively charged particles moving around it. However this leads to an

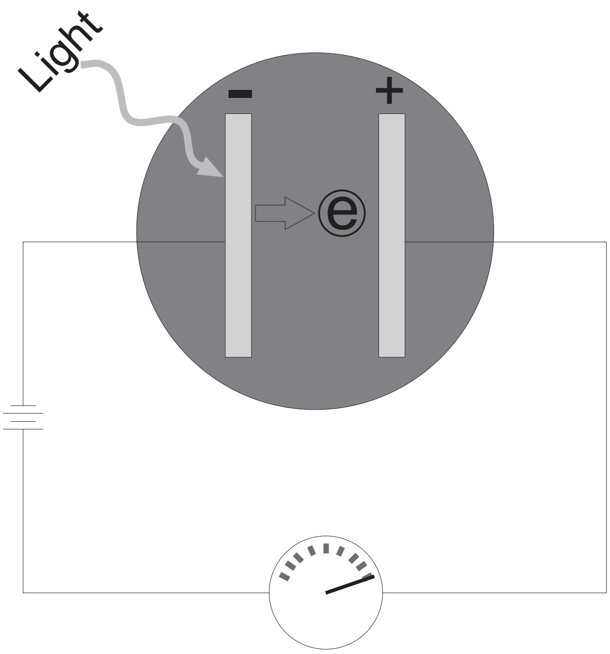


Figure 8.1: Photoelectric effect

inconsistency : An electron moving in a closed orbit at constant speed is actually accelerating since its velocity changes direction. Since an accelerating charge always radiates, the electron would continuously lose energy, spiralling. Hence, according to Newtonian physics, atoms are inherently unstable, a "prediction" that is contradicted by empirical facts.

In the 1900's, pioneers like Niels Bohr, Max Born, de Broglie and others, inspired by the quantum postulate of Planck, developed further **ad hoc** rules to explain the stability of the atom. Their model correctly predicted discrete energy levels in agreement with experimental data on spectral lines.

For our purposes it is sufficient to consider the crude Bohr Model in which the negatively charged electrons orbit a central positively charged nucleus at certain fixed radii: That is the radii of the orbit cannot be arbitrary but are fixed by some quantization rules. Each orbit can hold a certain maximum number of electrons. Atoms of different elements have different characteristic energy levels which are occupied by its electrons.

An electron moving in an allowed Bohr orbit does not radiate.

Electrons normally occupy the lowest energy states but can be excited to higher energy levels by an external input of energy. However this excited state is short lived and the electron returns to its stable ground state, emitting the excess energy in the form of a photon. The frequency f of the emitted radiation is related to the difference between the excited E_2 and

ground state energy level E_1, according to the relation

$$E_2 - E_1 = hf \tag{8.3}$$

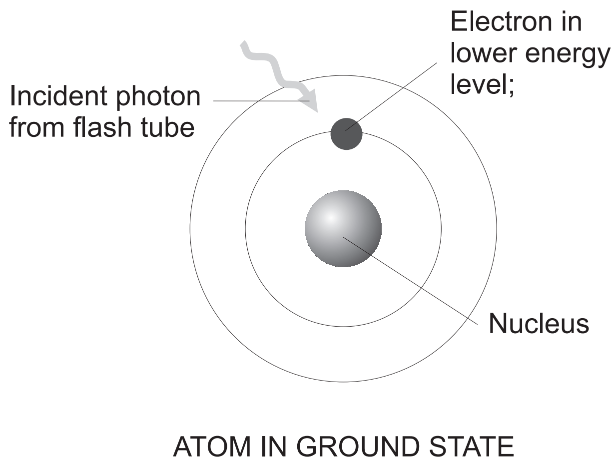


Figure 8.2: Absorption of energy by an atom

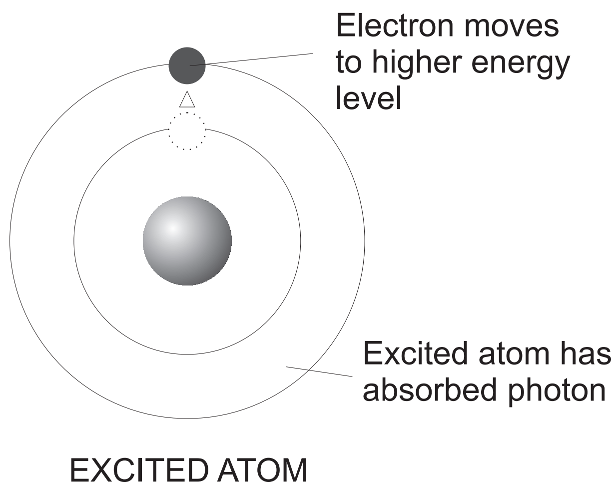


Figure 8.3: Atom in Excited State

8.1.3 The Electron as a Wave

In 1923, de Broglie conjectured that a particle might have a wave associated with it, just like an electromagnetic wave had particle charactersitics (the photon). For a particle of mass m moving at velocity v, de Broglie postulated

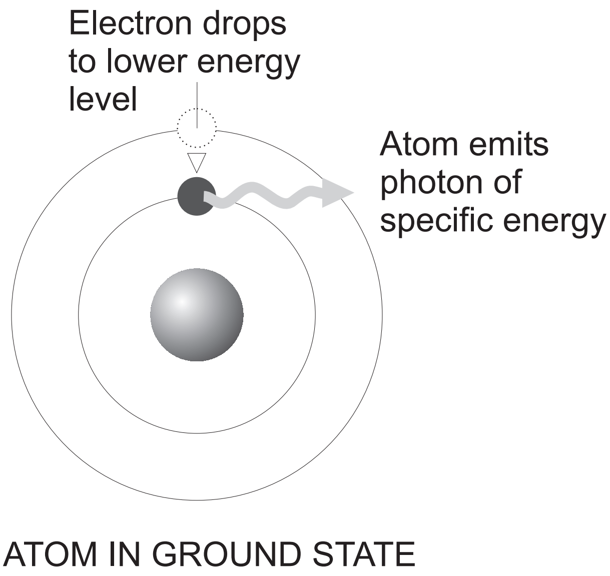


Figure 8.4: De-excitation of atom

an associated wave with a wavelength given by

$$\lambda = \frac{h}{mv} \tag{8.4}$$

The actual meaning of the wave was only understood later: The de Broglie wave not a physical wave, but instead, is a **probability wave**. Direct empirical evidence for the wave idea was also obtained: Davisson and Germer did an experiment in 1927 to verify the wave nature of electrons by scattering electrons off a crystal. The observed diffraction pattern confirmed theoretical predictions.

The Electron Microscope
In an optical microscope light scattered from a sample is magnified and then focussed onto the eye or a camera. The ability of a microscope to image an object clearly is limited by the wavelength used. Think of the analogy of water waves and how they are relatively unaffected by small stones in the pond — the wavelength must be of the order of the object for there to be significant change in the scattered wave. Hence optical microscopes are limited in their *resolving power*: They can only differentiate objects of the order of 10^{-6} m across. This is good enough to see some of the larger structures in a biological cell.
Better resolution demands the use of smaller wavelengths. In the 1930's Ernst Ruska decided to use electrons for this purpose. Recall that microscopic particles have wave-like properties, with a de Broglie wavelength inversely proportional to the particle's momentum. Electron microscopes can resolve objects up to $100,000$ times smaller than optical microscopes, thus reaching the atomic domain.
Unlike the glass lenses of optical microscopes, electron microscopes use magnetic fields to focus the beam.
The 1986 Nobel Prize was shared by Ruska (almost fifty years after his invention!) with the inventors of another powerful microscope mentioned later.

8.2 X-rays

X-rays have much shorter wavelength, or higher frequency, than even ultraviolet rays. Thus an X-ray photon carries much more energy than that of visible light or ultraviolet radition, and while the latter might be able to disrupt a single chemical bond, X-ray photon can destroy many bonds and cause permanent damage to molecules. Thus even low intensity (i.e small number of photons) X-rays are dangerous because they can cause significant permanent damage which becomes apparent only some time later.

The X-rays used for medical purposes are created by colliding fast-moving electrons with a heavy target. The rapid deceleration of the electrons produces high-energy electromagnetic rays in the X-ray range.

X-ray imaging makes use of the photoelectric effect: When an X-ray photon strikes an atom, it ejects an electron with a kinetic energy equal to the difference between the X-ray photon energy and the binding energy of the electron. Now it turns out that small atoms which have mostly weakly bound electrons, are quite unaffected by high energy X-rays while the tightly bound electrons in large atoms have a much greater probability of absorbibg the X-rays. Thus tissue which contains mostly small atoms like C,H,O,N, is relatively transparent to X-rays while bone with its larger Ca and P atoms

casts shadows in X-ray imaging.

8.3 Fluorescent Lamps

The inefficiencies of incandecent bulbs are solved by gas discharge lamps where instead of heating a filament, atoms of a gas are excited and used to give off the light. Fluorescent lamps are predominant in all modern lighting as they produce close to white light and do not waste much energy in producing heat or invisible infrared radiation.

A fluorescent lamp consists of a tube filled with a low-pressure (0.3% of atmospheric pressure) inert gas such as argon together with tiny amounts of mercury. There are electrodes at each end of the tube which inject electrons into the gas by thermal heating or other means when a current flows through them. The ejected electrons make the gas conducting and are accelerated by the potential difference between the electrodes, making collisons with the mercury atoms. The collisions often result in a transfer of energy from the electron to the mercury atom.

When a mercury atom receives some kinetic energy from a colliding electron, it uses it to push one of its own electrons into a higher energy state. When the atom relaxes back to its ground state, radiation is emitted. In fact, what is emitted by the mercury atoms is high energy ultraviolet light of wavelength 254 nanometres ($10^{-9}m$). Glass is opaque to this radiation however (and anyway as we have seen above the human eye does not register it). The trick to obtaining visible light is to coat the inside of the fluorescent tube with phosphor powder. Phosphor has the property of fluorescence, that is, it can convert ultraviolet light into visible light. Fluorescent lamps are between four to six times more efficient than incandescent bulbs.

8.4 Lasers

When light is emitted by an incandescent lamp, the photons from separate atoms move in different directions, have different wavelengths, different polarisations and also phases. Such light is called incoherent light.

If one could arrange for the individual photons to be emitted in the same direction, with the same wavelength, polarisation and phase, then the resultant light would be more intense and focussed. Such coherent light is what lasers produce.

In a laser, a single photon is used to stimulate other atoms in the laser medium to produce identical copies of that one photon. Thus through the amplification process an intense coherent beam is produced.

Technically, the atoms in the laser medium are first pumped to an excited state by an external energy source such as currents as in the case of diode lasers. When a seed photon (produced for example by spontaneous emission

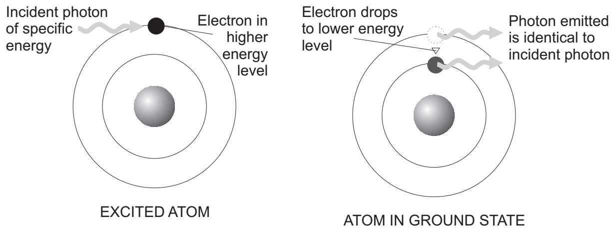


Figure 8.5: Stimulated Emission

by one of the atoms) passes close by those excited atoms, the electric field of the photon stimulates the atoms to return to a lower energy state by releasing photons identical to the seed photon. The laser medium is placed between two mirrors so that the process can be repeated many times. One of the mirrors is actually semi-transparent so that the amplified light beam can leave it.

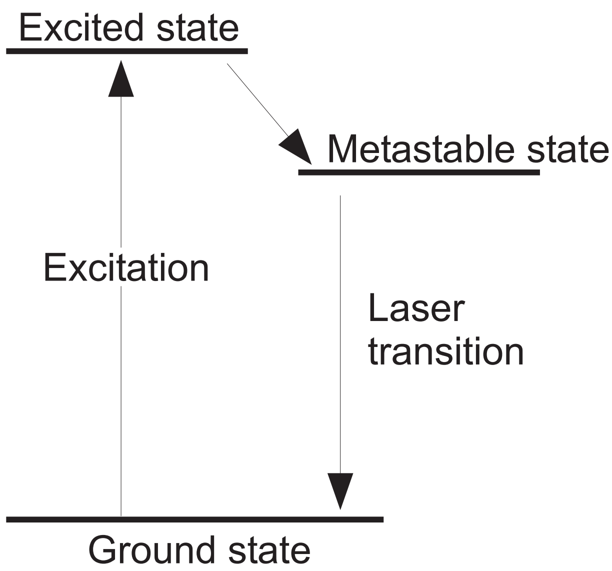


Figure 8.6: Energy States

Diode lasers are found in laser pointers and CD players.

8.4.1 Optical Fibers

We mentioned earlier how coaxial cables can be used to transmit electromagnetic waves over long distances with little loss. An optical fiber does the same for light. A thin solid core of very pure high refcative index glass is surrounded by a thicker layer of low-refractive index glass. Light travels

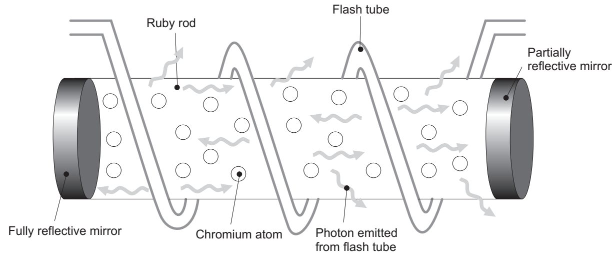


Figure 8.7: Ruby laser

along the core, no matter how the fiber is bent, because of the process of total internal reflection.

A 1550nm diode laser is used to generate short pulses of light which carry information for optical communication. The much higher frequency of light compared to radio waves implies that much more information can be transmitted in the visible spectrum.

8.5 Light Emitting Diodes

Unlike incandescent light bulbs, LED's do not produce light by heating a filament but rather directly by electrons moving from one energy level to another in a semiconductor. Thus LED's are more efficient (in energy conversion terms) and reliable (they dont have a filament that can burn out) and are beginning to replace light bulbs in many applications.

As its name suggests, a semiconductor is not a very good conductor of electricity. Consider pure silicon: Each atom has four electrons in its outermost shell and thus it can bond (covalent bonds) with four other silicon atoms to have a shared total of eight electrons that completes its shell. In this way pure silicon is a lattice of perfectly bonded silicon atoms. If some energy is provided, then some electrons will escape from the bonds, leave the atoms and become available to conduct electricity. However the electrical properties of semiconductors becomes significant only when they have been **doped** with some impurities, meaning, some of the silicon atoms have been replaced with other atoms such as arsenic or gallium.

If arsenic or phosphorus are used, then one gets an **N-type** semiconductor. The name comes about because arsenic and phosphorus have five electrons in their outer shell and so when those atoms fit into the silicon lattice free electrons are made available: Electrons are negatively (N) charged.

On the other hand boron or gallium atoms each have only three outer electrons, and so when they are used a deficit of electrons occurs in the lattice. A "missing" electron behave equivalently to a positive (P) charge or **hole**. These holes are the charge carriers of the **P-type** semiconductors.

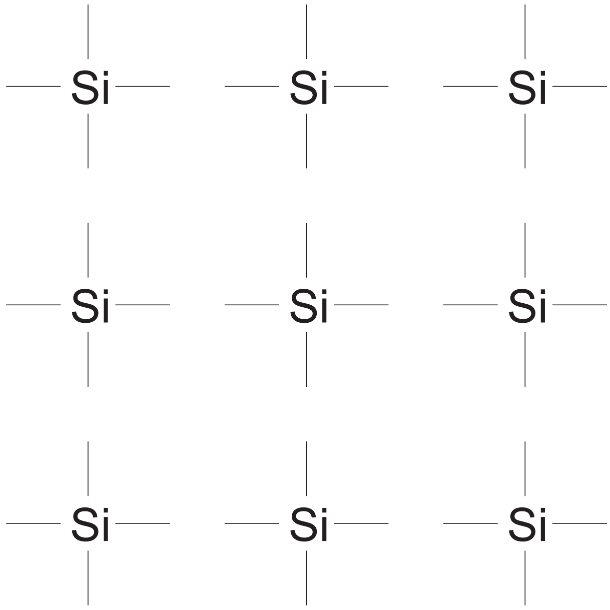


Figure 8.8: Pure silicon

A semiconductor **diode** is formed when an N-type semiconductor is placed adjacent to a P-type semiconductor. At the boundary naturally the free electrons from one will combine with the holes of the other to form a **depletion zone** which is now insulating. Note that as the depletion zone forms, an electric field is created: The electric field points from the N zone to the P zone and is a result of the unbalanced charges on the atoms in the depletion zone.

If electrodes are attached to the free ends of the diode and an external voltage (exceding a certain minimum strength) supplied, current will flow when the N zone is connected to the negative end of the battery and the P end to the positive end, and not for opposite polarities (whereby the depletion zone widens).

When the diode is in its conducting state, electrons from the N side fall into holes on the P side and emit photons. For silicon diodes the energy is in the infrared range. To obtain visible light a material such as aluminium-gallium-arsenide has to be used as this is has a larger energy gap between its conduction and lower energy bands. Practical LED's are designed so that a substantial number of the photons is emitted outward in a particular direction

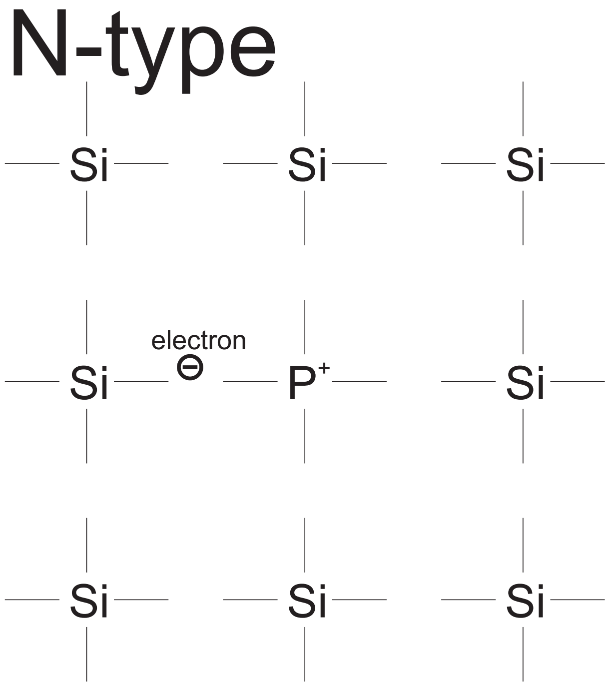


Figure 8.9: N-type semiconductor

8.6 Quantum Mechanics

While the work of Planck, Einstein, de Broglie, Bohr, Born and others started the quantum theory, the initial framework, as briefly discussed above, consisted of an *ad hoc* mixture of classical concepts and some new rules, and it was only much later that a fully consistent **Quantum Mechanics**, replacing classical mechanics, was developed. Of course classical mechanics is recovered as an approximation valid in our macroscopic world.

Unlike classical mechanics which deals with well-defined particle trajectories, quantum mechanics deals with probability amplitudes.

Although the ordinary world we perceive by our physical senses leads to Newtonian physics, reality as described by quantum mechanics is counter-intuitive. The student can take some consolation from the fact that even professional physicists have a hard time "visualising" (it is doubtful if anyone can) the bizarre quantum world. As stated above, this is not surprising since human senses have developed to cope with a macroscopic classical environment and all our mental pictures and language reflect our experience in the macroscopic domain. Some authors even prefer to use new words such as "**wavicle**" in qualitative discussions of the dual wave/particle nature of light and matter. Nevertheless, it must be emphasized that **quantum mechanics provides a well-defined and rigorous mathematical framework**,

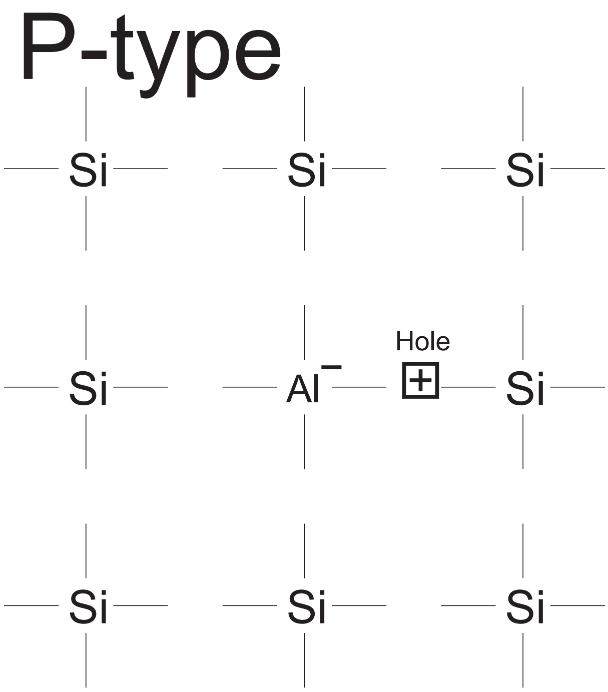


Figure 8.10: P-type semiconductor

using admittedly abstract and non-classical concepts, which allows one to calculate measurable quantities.

We can extend what we perceive with our five senses with experimental devices and instruments to probe more deeply into nature's secrets, where the quantum aspect of nature become more apparent. A hundred years of comparing theory with experiment has shown that quantum mechanics is correct even if it is highly non-intuitive.

The following summarises some of the unusual features of quantum theory.

1. **Energy can come in discrete packets called quanta**. For example, a quantised excitation of the elctromagnetic wave (photon) behaves very much like a particle; The converse is also true, that particles can behave like a wave — a probability wave. (In quantum mechanics one solves the Schroedinger equation, $i\hbar\frac{\partial\psi}{\partial t} = H\psi$, to determine the wavefunction ψ which is the probability wave. Whereas in classical mechanics one uses Newton's Second law to determine the trajectory of a particle).

2. In quantum mechanics, since the fundamental description is in terms of probability waves, a partice can exist in the superposition of two

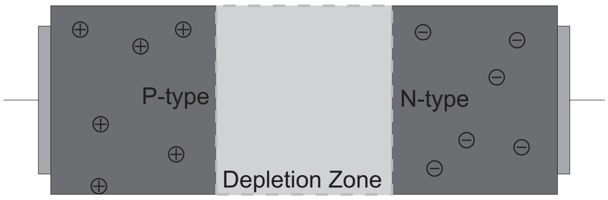


Figure 8.11: A semiconductor diode

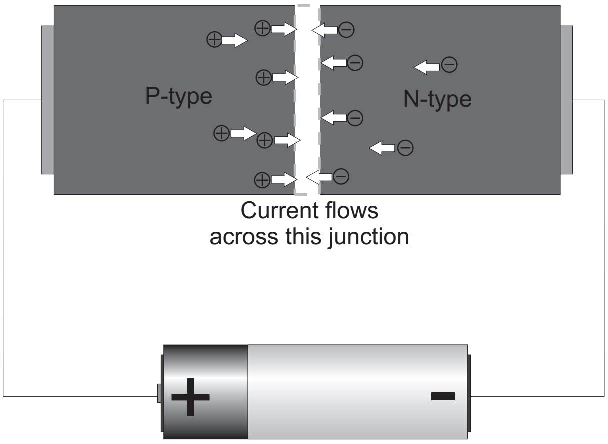


Figure 8.12: Diode in Conducting State

distinct (from the classical viewpoint) states. For example, an electron can exist as a superposition of spin-up and spin-down states. A single (eigen)state is slected only when a measurement is made.

3. Quantum Mechanics predicts that it is possible for particles to move into regions of space which are classically forbidden on energetic grounds. This gives rise to the phenomenon of tunneling. The phenomenon of tunneling is not just a curiosity but is actually required to explain physical processes such as alpha-decay, and is also used explicitly in the design of some modern microelectronic devices.

4. An intrinsic indeterminacy enters the description of quantum particles. This is encoded in the famous Heisenberg uncertianty principle

$$\Delta x \Delta p \geq \frac{\hbar}{2}, \tag{8.5}$$

which states that the more precisely we try to localise (x) a particle in a particular direction, the more uncertain we become of its momentum (p) in that direction, and *vice versa*

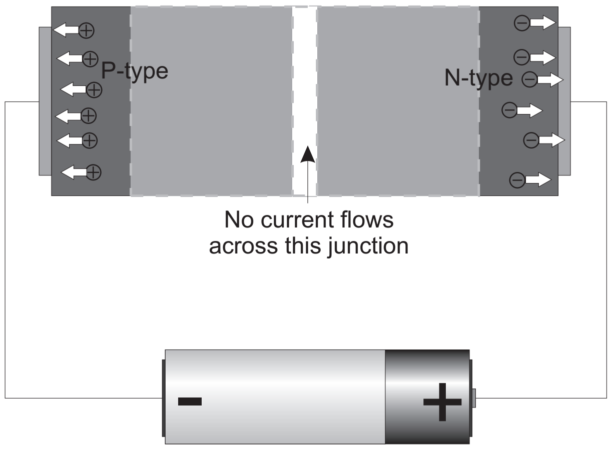


Figure 8.13: Diode in Insulating State

the more precisely we try to determine the momentum of a particle along a particular direction, the more uncertain we become of its location along that direction.

It is very important to note that the uncertainties discussed above are NOT due to imperfect apparatus but are intrinsic to the quantum nature of the particle: It is necessarily described by an extended wavefunction and hence probabilities rather than precise postion and velocities which would have given rise to precise classical trajectories (but an eventual disagreement with experimental data!)

The Scanning Tunneling Microscope
In the 1970's G. Binnig and H. Rohrer used the idea that electrons can tunnel through classically forbidden regions to construct the Scanning Tunneling Microscope. Instead of scattering electrons off the surface to be examined, a very sharp probe (whoe tip is just a single atom) is suspended just a few atomic radii above the surface of a sample. The tunneling current from the sample is measured and by moving the tip across the surface, the landscape with its peaks and valleys can be imaged. In this way images of single atoms have been obtained. The STM can also be used to drag and place atoms at desired locations. Binnig and Rohrer shared the 1986 Nobel Prize with E. Ruska, the inventor of the electron microscope.

8.7 Summary

Quantum mechanics explains not only the structure and stability of matter at microscopic scales, but also phenomena such as super-conductivity and super-fluidity which are more macroscopic manifestations of quantum behaviour. Applied science from that of semiconductors to more complex technologies such as MRI (Magnetic Resonance Imaging), lasers, the design of new drugs and modern materials science all draw on the principles of quantum theory.

Finally, we mention three areas of research where frontier applications of quantum mechanics could potentially have a dramatic impact on life and technology in the 21st century: Quantum computers, quantum cryptography and nanotechnology. Nanotechnology is briefly described in the next subsection while in the references the interested reader will find some articles on quantum cryptography and computers.

8.7.1 Nanotechnology

Nanotechnology is the broad name scientists have given to a rapidly evolving field that its promoters promise will be the revolution of the 21st century. Nano is a scientific prefix meaning one-billionth. So a nanometre is one-billionth of a meter, i.e. about 10,000 times smaller than the diameter of a human hair! Nanotechnology concerns the study and design of structures in the 1-100 nanometre range. New materials and devices will be created at this scale not only by the usual chemical methods, but also either by directly (physically) putting them together molecule by molecule, or by imitating biology and using a procedure of self-assembly.

Nanotechnology did not take off immediately after Feynmans prophetic lecture as there seemed to be little to motivate researchers. Its potential was independently recognized and propagated many years later by Eric Drexler who envisaged a futuristic world of tiny robots creating all that we could desire by building things up atom by atom. For example, in Drexlers futruristic world one would dump raw chemicals into an assembler, key in a few numbers (almost like operating a microwave), and voila a few minutes later one would have fried chicken or pasta as chosen!

A typical atom has a diameter of one-tenth of a nanometre, **so nanostructures consist of only a few atoms, thus making their properties quite different from larger bulk objects**. In the beginning most scientists were skeptical about directly manipulating such small objects and whether the endeavour would eventually be useful. However as Feynman and Drexler realized, Nature provides abundant proof that functional and useful nanometre scale structures can be formed: The components of a typical human cell, such as the DNA, membranes and enzymes are some examples. Furthermore, genetic engineering is an example of a mature field that

provides evidence that nanoscale manipulation is possible. In essence then nanotechnology will attempt to reproduce in the inorganic world what has been so successful in biology.

The original focus of nanotechnology was the design of miniature machines and robots, and some of these nanomachines have actually been created. More generally, nanotechnology nowadays includes several other sub-disciplines: the creation of new materials with novel mechanical properties, such as carbon nanotubes and aluminium composites; nanoelectronics (quantum wells and wires); nanobiotechnology; and nanomedicine.

It is generally believed that in the coming decades nanotechnology will have a dramatic impact in diverse areas such as computing, communication, energy, environment, medicine and manufacturing. Some of the possibilities that are being considered are: 'smart' molecular devices for the treatment of diseases, construction of super-efficient solar cells from nanoengineered materials, smaller and faster computers, and biosensors for environmental monitoring and cleansing.

A feature of nanotechnology research is its intrinsically interdisciplinary nature involving the fundamental sciences, computer modeling and engineering. There are thus likely to be exciting developments and prospects in this frontier field. However it is important to distinguish actual research and progress in the field from some of the popular science hype that it has attracted. A search on the Internet reveals that nano is now prefixed to a host of terms (and even used by itself!) regardless of whether it has anything at all to do with things at the nanometer scale or nanotechnology. Historically, similar problems have afflicted other words such as quantum, relativity, and chaos. Using other new-age fields such as genomics and proteomics as guides, it is likely that significant progress in nanotechnology will take time but that the payoffs are potentially enormous.

8.8 Curiosity: Can The Human Eye Detect a Single Photon of Light?

There are two types of receptors on the retina at the back of the human eye. Colour vision is possible because of the cones while vision in low intensity light is possible because of the rods. In an experiment performed in 1942, it was established statistically that a individual rod was sensitive to a single photon! *However this does not mean that we can conciously detect a single photon.* In fact if our eyes were that sensitive, our brain would be swamped with noisy stimuli. Indeed, we have evolved to the stage where at least 5 (to 9) photons must reach the rods within a time period of 100ms before filters will activate a signal to the brain and for us to be conscious of seeing something. It is also important to note that only about 10% of the light that reaches the eye actually is received by the rods: 3% is re-

flected from the cornea, 47% is absorbed by various nonsensing parts and according to Ref.[**??**] an incredible 40% actually falls between the rods and goes undetected. From the above numbers one deduces that between 500 to 900 photons must arrive at the eye every second for our brain to register a conscious signal.

8.9 Exercises

1. (a) Give an example of an experiment that illustrates the wave nature of light.
   (b)Give an example that illustrates the particle nature of light.
   (c) So is light a wave or a particle ?
   (d) If a beam of yellow light has exactly the same energy as a beam of green light, which beam has more photons?

2. (a) What evidence is there for quantised energy states?
   (b) It is sometimes stated that everything is quantised at the microscopic level. Is this true? Why?
   (c) It is sometimes stated that light energy is quantised. Is this true? Clarify.

3. (a) What is the difference between a planet moving around the sun and the motion of an electron around a nucleus?
   (b) Is the electron wave a physical wave or a probability amplitude? Explain the difference.

4. What is the difference between
   (a) An ordinary light microscope and an electron microscope,
   (b) An electron microscope and a scanning tunneling microscope.
   (c) Why is the electron microscope used in imprinting patterns on semiconductor chips? Can the scanning tunneling microscope be used for similar purposes?

5. (a) Calculate your de Broglie wavelength at your walking speed.
   (b) Estimate the speed of electrons used in an electron microscope.

6. (a) How are CAT scans different from X-ray imaging ? (Hint: See Ref.[2]).
   (b) Compare the energy of one photon of X-rays with that of UV light.

7. (a) Why are fluorescent tubes usually cylindrical ?
   (b)What are the hazards involved in the disposal of fluorescent tubes ?

8. In contrast to incandescent bulbs and fluorescent lamps, light sticks use no electric power source. How do they produce light? (Hint: see Ref.[5])

9. "Black light" are used in several places to reveal invisible images or writing. How do they work ?

10. Some detergents claim to produce "whiter than white" clothes on washing. How do they work?

11. (a) Where does the energy of an intense laser beam come from ?
(b) What are some of the applications of lasers ?
(c) What different types of lasing mediums are there ?
(d) What kind of laser would you use to cut through steel ?
(e) Are supermarket laser scanners and laser pointers dangerous in any way ?
(Hint: See Ref.[7]).

12. (a) What are the advantages of optical fibers compared to conventional copper wires for communications ?
(b) What are some other applications of optical fibers?
(Hint: See Ref.[8]).

13. If LED's are so wonderful, why has their use been so slow to spread ? Where would their use be ideal even with the current disadvantages?

14. Optional: X-rays are not suitable for imaging tissue. An alternative technique which allows this to be done is called Magnetic Resonance Imaging (MRI). You can read about this in Ref.[3].

15. Optional: The process by which alpha radiation is produced was explained by Gamov in terms of tunneling of particles through the nucleus.
(a) Can you tunnel through a wall? Why and How ?
(b) What is a tunnel diode?

16. Optional: Discuss some of the likely benefits of nanotechnology. How many of these have been realised ?

8.10 References

1. Tunneling at
http://phys.educ.ksu.edu/vqm/html/qtunneling.html

2. CAT scans at http://www.howstuffworks.com/cat-scan.htm

3. MRI at http://www.howstuffworks.com/mri.htm

4. How Fluorescent Lamps Work at http://www.howstuffworks.com/fluorescent-lamp2.htm

5. Light Sticks at http://www.howstuffworks.com/light-stick.htm

6. Black Light at http://www.howstuffworks.com/question59.htm

7. Lasers at http://www.howstuffworks.com/laser.htm

8. Fiber Optics at http://www.howstuffworks.com/fiber-optic.htm

9. Nanotechnology at Sandia National Laboratory : http://nano.sandia.gov/

10. Quantum Computers at
http://www.howstuffworks.com/quantum-computer.htm
and http://www.cs.caltech.edu/~westside/quantum-intro.html and
http://www.qubit.org/oldsite/intros/comp/comp.html

11. Quantum Cryptography at
http://www.dhushara.com/book/quantcos/aq/qcrypt.htm#anchor618596
and http://www.qubit.org/library/intros/crypt.html
and http://www.cyberbeach.net/~jdwyer/quantum_crypto/quantum1.htm

12. The Sensitivity of the Human Eye at
http://www.public.iastate.edu/~physics/sci.physics/faq/see_a_photon.html
and
http://arapaho.nsuok.edu/~salmonto/VSII/Lecture1.pdf